

STIC-Biotech/ChemLib *CRFE*

145402 *mg*

From: Swope, Sheridan
Sent: *79280* Wednesday, February 16, 2005 8:59 PM
To: STIC-Biotech/ChemLib
Subject: 10/701,200

For 10/701,200 pls search:
SID 6 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

aa - 437

Beverly Shears

*copy
this page*

mg

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: *2/19/05*
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search
NA Sequence: # _____
AA Sequence :# _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

Swope, S.
10/701200 Page 1
Seq. ID 6 w/ Interf

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - protein search, using BW model

Run on: February 24, 2005, 14:00:08 / Search time 166 Seconds

(without alignments)
1018.158 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTWPHYLADIRFCWMP.....FNSMLSEIGPKGKREVSH 437

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_16Dec04.*

- 1: Geneseqp19808.*
- 2: Geneseqp19808.*
- 3: Geneseqp20008.*
- 4: Geneseqp20018.*
- 5: Geneseqp20028.*
- 6: Geneseqp20038.*
- 7: Geneseqp20038.*
- 8: Geneseqp20048.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 2284 | 100.0 | 437 | 5 | ABB09168 |
| 2 | 2284 | 100.0 | 437 | 5 | AAE22299 |
| 3 | 2284 | 100.0 | 437 | 5 | ABG61553 |
| 4 | 1442 | 63.1 | 410 | 6 | AAU45865 |
| 5 | 1442 | 63.1 | 410 | 6 | ABM42384 |
| 6 | 659 | 28.9 | 184 | 8 | ADS25315 |
| 7 | 306.5 | 13.4 | 350 | 8 | ADN25455 |
| 8 | 282 | 12.3 | 488 | 5 | ABB93840 |
| 9 | 281.5 | 12.3 | 485 | 3 | AAG39480 |
| 10 | 270 | 11.8 | 448 | 6 | ABU19327 |
| 11 | 270 | 11.8 | 483 | 2 | AAK71580 |
| 12 | 265 | 11.8 | 469 | 2 | AAK71581 |
| 13 | 263 | 11.5 | 426 | 2 | AAK71583 |
| 14 | 262 | 11.5 | 485 | 2 | AAK71579 |
| 15 | 261.5 | 11.4 | 496 | 7 | ADC07918 |
| 16 | 261 | 11.4 | 345 | 8 | ADS30228 |
| 17 | 257.5 | 11.3 | 500 | 5 | ABB93109 |
| 18 | 253 | 11.1 | 319 | 8 | ADN20252 |
| 19 | 251 | 11.0 | 473 | 3 | AAK36438 |
| 20 | 251 | 11.0 | 473 | 5 | ABB93145 |
| 21 | 251 | 11.0 | 537 | 3 | AAK36437 |
| 22 | 250.5 | 11.0 | 357 | 8 | ADN27016 |
| 23 | 247 | 10.8 | 365 | 3 | AAK39481 |
| 24 | 246.5 | 10.8 | 987 | 6 | ABK52589 |
| 25 | 246.5 | 10.8 | 987 | 7 | ADK61892 |

| | | | | | | |
|----|-------|------|-----|---|----------|--------------------|
| 26 | 246.5 | 10.8 | 987 | 8 | ADS43568 | Ad43568 Bacterial |
| 27 | 243 | 10.6 | 462 | 5 | ABB93203 | Ab93203 Herbicida |
| 28 | 240.5 | 10.5 | 320 | 4 | AAU38192 | Aau38192 Salmonell |
| 29 | 240.5 | 10.5 | 320 | 6 | ABU48101 | Abu48101 Protein e |
| 30 | 240.5 | 10.5 | 320 | 6 | ABU47109 | Abu47109 Protein e |
| 31 | 240.5 | 10.5 | 320 | 6 | ABU45064 | Abu45064 Protein e |
| 32 | 239.5 | 10.5 | 318 | 8 | ADN26999 | Adn26999 Bacterial |
| 33 | 239 | 10.5 | 462 | 3 | AAK28969 | Aak28969 Arabidops |
| 34 | 238.5 | 10.4 | 489 | 3 | AAK34829 | Aak34829 E. coli c |
| 35 | 238.5 | 10.4 | 320 | 4 | AAU34829 | Aau34829 Protein e |
| 36 | 238.5 | 10.4 | 320 | 6 | ADS45208 | Ad45208 Bacterial |
| 37 | 238.5 | 10.4 | 320 | 6 | ABU28847 | Abu28847 Protein e |
| 38 | 234.5 | 10.3 | 327 | 6 | ABU50166 | Abu50166 Protein e |
| 39 | 234 | 10.2 | 509 | 3 | AAK47222 | Aak47222 Arabidops |
| 40 | 234 | 10.2 | 525 | 5 | ABB93912 | Ab93912 Herbicida |
| 41 | 232 | 10.2 | 461 | 6 | ABU48474 | Abu48474 Protein e |
| 42 | 232 | 10.2 | 462 | 4 | AAU42927 | Aau42927 Propionib |
| 43 | 232 | 10.2 | 462 | 6 | ABM39446 | Abm39446 Propionib |
| 44 | 229.5 | 10.0 | 295 | 8 | ADN26202 | Adn26202 Bacterial |
| 45 | 229.5 | 10.0 | 320 | 6 | ABU28456 | Abu28456 Protein e |

ALIGNMENTS

| | | |
|----------|---|----------------------------|
| RESULT 1 | ABB09168 | standard; protein; 437 AA. |
| ID | ABB09168 | |
| XX | ABB09168; | |
| DT | 01-JUL-2002 | (first entry) |
| XX | | |
| DE | Methylomonas pyrophosphate dependent phosphofructokinase SEQ ID NO:16. | |
| XX | | |
| KW | Methylomonas; methanotrophic; carbon metabolism; carbon flux pathway; | |
| KW | transaldolase; fructose biphosphate aldolase; KHG/KDPG aldolase; food; | |
| KW | pyrophosphate dependent phosphofructokinase; phosphoglucutase; feed; | |
| KW | glucose 6 phosphate isomerase; 6-phosphogluconate dehydratase; methane; | |
| KW | single carbon substrate; 1 dehydrogenase; enzyme; methanotroph; methanol; | |
| KW | isoprenoid; carotenoid pigment. | |
| XX | | |
| OS | Methylomonas sp. | |
| XX | | |
| PN | WO200220796-A2. | |
| PD | 14-MAR-2002. | |
| XX | | |
| PF | 28-AUG-2001; 2001WO-US026730. | |
| XX | | |
| PR | 01-SEP-2000; 2000US-0229906P. | |
| XX | | |
| PA | (DUPO) DU PONT DE NEMOURS & CO E. I. | |
| XX | | |
| PI | Koffas M, Odom JM, Norton KC, Ye RW; | |
| XX | | |
| DR | WPI: 2002-362250/39. | |
| DR | N-PSDB; ABI51513. | |
| XX | | |
| PT | New polynucleotide encoding a Methylomonas sp. carbon flux enzyme useful | |
| PT | for altering carbon flow through methanotrophic bacteria, utilized for | |
| PT | production of single cell protein and commercially valuable | |
| PT | polysaccharides. | |
| XX | | |
| PS | Claim 4; Page 66-68; 73pp; English. | |
| XX | | |
| CC | The present invention describes isolated polynucleotides (I) encoding a | |
| CC | Methylomonas sp. carbon flux enzyme, consisting of: transaldolase; | |
| CC | fructose biphosphate aldolase; KHG/KDPG aldolase; phosphoglucutase; | |
| CC | pyrophosphate dependent phosphofructokinase; glucose 6 phosphate | |
| CC | isomerase; 6-phosphogluconate dehydratase; and glucose 6 phosphate 1 | |
| CC | dehydrogenase enzymes. (I) overexpression is useful for altering carbon | |

CC flow through a methanotrophic bacteria, where the bacteria grows on a C1
CC carbon substrate of methane and methanol, and comprises a functional
CC Embden-Meyerhof carbon pathway comprising a gene encoding a pyrophosphate
CC dependent phosphofructokinase enzyme, where the bacteria is preferably
CC Methylobionas 16a American Type Culture Collection (ATCC) PTA 2402, and
CC where the carbon flux gene is: (i) over-expressed on a multicopy plasmid;
CC (ii) operably linked to an inducible or regulated promoter; (iii)
CC expressed in an antisense orientation; or (iv) disrupted by insertion of
CC foreign DNA into the coding region. The manipulated methanotrophs are
CC useful for the energy efficient conversion of single carbon substrates
CC such as methane and methanol to commercially useful products in the food,
CC feed and materials industries, and preferably for the production of
CC single cell protein, and for producing polysaccharides, used as
CC thickeners in food and non-food industries, and isoprenoid compounds and
CC carotenoid pigments of various carbon lengths. The present sequence
CC represents a Methylobionas pyrophosphate dependent phosphofructokinase
CC from the present invention
CC
XX
SQ Sequence 437 AA:
Query Match 100.0%; Score 2284; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 6e-210;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DVTWPHYLTADIRFCHEPFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWPHYLTADIRFCHEPFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEITCYRGYKGLLGDSYPTAEVRKKAAGVLOFGSGSVIGNSVVKLTNVDCYKR 120
DB 61 IDPSIEITCYRGYKGLLGDSYPTAEVRKKAAGVLOFGSGSVIGNSVVKLTNVDCYKR 120
QY 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAFLARNNGYGLTVIGLPTV 180
DB 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAFLARNNGYGLTVIGLPTV 180
QY 181 DNDVFPITKQSIGAATAEQGARFYFNVAENNANPRMLIVHEWGRNCGWLTAAATQRYR 240
DB 181 DNDVFPITKQSIGAATAEQGARFYFNVAENNANPRMLIVHEWGRNCGWLTAAATQRYR 240
QY 241 KLDRAEWLPBGLTRESYEYHAFVPEMAIDLEAERLREVMKVCVNIFFVSEGAGV 300
DB 241 KLDRAEWLPBGLTRESYEYHAFVPEMAIDLEAERLREVMKVCVNIFFVSEGAGV 300
QY 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKMFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
DB 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKMFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVLAIEFPRIKGGKPFNIDTWENS 420
DB 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVLAIEFPRIKGGKPFNIDTWENS 420
QY 421 MLSEIGQPKGKVEYSH 437
DB 421 MLSEIGQPKGKVEYSH 437
RESULT 2
ID AAE22299 standard; protein; 437 AA.
XX AAE22299,
AC AAE22299,
XX
DT 07-AUG-2003 (revised)
DT 25-JUL-2002 (first entry)
XX
DB Methylobionas 16a sp. phosphofructokinase pyrophosphate dependent enzyme.
XX Carotenoid; isoprenyl pyrophosphate; anthraxanthin; aetaxanthin; diet;
KW anti-oxidant; steroid; flavour; fragrance; electro-optic application;
KW aquaculture; enzyme; phosphofructokinase pyrophosphate.
XX
OS Methylobionas sp.

XX
FN W0200218617-A2.
XX
XX 07-MAR-2002.
XX
PD
XX
PE 04-SEP-2001; 2001WO-US027420.
XX
XX
PR 01-SEP-2000; 2000US-0229858P.
PR 01-SEP-2000; 2000US-0229907P.
XX
XX
PA (DUPO) DU PONT DE NEMOURS & CO E. I.
PI Broscowicz PC, Cheng Q, Dicosimo DJ, Koffas M, Miller ES;
PI Odom JM, Piatagallo SK, Rouviere FB;
DR N-PSDB; AAD35497.
XX
XX
PT Producing carotenoid compounds e.g. anthraxanthin and aetaxanthin, by
PT using microorganisms having a nucleic acid molecule encoding enzymes in
PT the carotenoid biosynthetic pathway and which metabolize single carbon
PT substrates.
PS
PS Claim 8; Page 104-106; 156pp; English.
XX
XX The invention relates to a method for producing carotenoid compounds. The
XX method comprises a transformed metabolising host cell, comprising
XX suitable levels of isoprenyl pyrophosphate and a nucleic acid molecule
XX encoding an enzyme in the carotenoid biosynthetic pathway, under the
XX control of regulatory sequences, and contacting the host cell with carbon
XX substrate to produce a carotenoid compound. The method is useful for
XX producing carotenoid compounds such as anthraxanthin and aetaxanthin, by
XX using microorganisms having a nucleic acid molecule encoding enzymes in
XX the carotenoid biosynthetic pathway and which metabolize single carbon
XX substrates. The carotenoids have potent anti-oxidant properties useful in
XX diet, and aquaculture elements. The carotenoids are also useful as
XX intermediates in the synthesis of steroids, flavours and fragrances and
XX compounds for potential electro-optic applications. The present sequence
XX is Methylobionas 16a sp. phosphofructokinase pyrophosphate dependent
XX enzyme used in the invention. (Updated on 07-AUG-2003 to correct OS
XX field.)
SQ Sequence 437 AA:
Query Match 100.0%; Score 2284; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 6e-210;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DVTWPHYLTADIRFCHEPFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWPHYLTADIRFCHEPFLNPNFYTLNPKKVAITTAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEITCYRGYKGLLGDSYPTAEVRKKAAGVLOFGSGSVIGNSVVKLTNVDCYKR 120
DB 61 IDPSIEITCYRGYKGLLGDSYPTAEVRKKAAGVLOFGSGSVIGNSVVKLTNVDCYKR 120
QY 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAFLARNNGYGLTVIGLPTV 180
DB 121 GLVKEGEDPQKVAADQVLDKGDVILHTIGDDTNTAAADLAFLARNNGYGLTVIGLPTV 180
QY 181 DNDVFPITKQSIGAATAEQGARFYFNVAENNANPRMLIVHEWGRNCGWLTAAATQRYR 240
DB 181 DNDVFPITKQSIGAATAEQGARFYFNVAENNANPRMLIVHEWGRNCGWLTAAATQRYR 240
QY 241 KLDRAEWLPBGLTRESYEYHAFVPEMAIDLEAERLREVMKVCVNIFFVSEGAGV 300
DB 241 KLDRAEWLPBGLTRESYEYHAFVPEMAIDLEAERLREVMKVCVNIFFVSEGAGV 300
QY 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKMFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
DB 301 EAIYAEMQAKQEVPRDAFGHIKLDVAVNPKMFGSQFQOMIGAEXTLVOKSGYFAPASAS 360
QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVLAIEFPRIKGGKPFNIDTWENS 420

|||||
Db 361 NVDDMRLLKSCADLAVECAFRESGVIGHDEBNGNVLAIEFPRIKGGKPFNIDTDFNS 420
Qy 421 MLSEIGOPKGGKVEVSH 437
Db 421 MLSEIGOPKGGKVEVSH 437
RESULT 3
ABG61553
ID ABG61553 standard; protein; 437 AA.
XX ABG61553;
AC
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX High growth methanotrophic bacterial strain polypeptide #3.
DE
XX
XX High growth methanotrophic bacterial strain; C1 carbon substrate; enzyme;
KM methane; methanol; Embden-Meyerhof carbon flux pathway; 16S RNA;
KM pyrophosphate dependent phosphofructokinase; nitrogen-containing compound;
KM ammonia; nitrate; nitrite; nitrogen; pigment; oxygen; landfill;
KM methane-containing environment; waste water treatment system; isoprenoid;
KM nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide.
XX
XX Methylomonas sp.
OS
XX
XX WO200220728-A2.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 28-AUG-2001; 2001WO-US026827.
PF
XX
XX 01-SEP-2000; 2000US-0229858P.
PR
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA
XX
XX Koffas M, Odom JM, Schenzle A;
PI
XX
XX WPI: 2002-454358/48.
DR
XX
XX N-PSDB; ABR63332.
DR
XX
XX New high growth methanotrophic bacterial strain, useful for producing
PT single cell proteins, grows on a C1 carbon substrate, and comprises a
PT functional gene encoding in Embden-Meyerhof carbon pathway.
XX
XX
PS Claim 7, Page 85-87; 157pp; English.
XX
XX The invention relates to a high growth methanotrophic bacterial strain,
XX which grows on a C1 carbon substrate e.g. methane and methanol, and
XX comprises a functional Embden-Meyerhof carbon flux pathway comprising a
XX gene coding a pyrophosphate dependent phosphofructokinase enzyme or a 16S
XX RNA. The bacterial strain is useful for the production of single cell
XX protein and for the biotransformation of a nitrogen-containing compound,
XX e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the
XX production of a feed product comprising a protein, carbohydrates and a
XX pigment and for reducing oxygen demand, for removing nitrates and
XX nitrates in methane-containing environments such as landfill, waste
XX water treatment systems or anywhere that methane, oxygen and nitrates are
XX present. The bacterial strain of the invention can be used as a
XX denitrifying agent for the conversion of nitrate or nitrite to nitrous
XX oxide with methane or methanol as a carbon source. It is also used in the
XX production of biomass including proteins, carbohydrates and a wide
XX variety of pigments (particularly for isoprenoid pigments for the purpose
XX of generating animal feeds), in production of terpenoid and carotenoid
XX compounds, useful as pigments and as monomers in polymeric materials and
XX in production of exopolysaccharides at high levels. Sequences ABG61551-
XX ABG61590 represent high growth methanotrophic bacterial strain proteins
XX of the invention
XX
XX Sequence 437 AA;
SQ
Query Match 100.0%; Score 2284; DB 5; Length 437;

Best Local Similarity 100.0%; Pred. No. 6e-210;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DVVTWPHLTADIRFCHEPFLNFNFYTLNNKRRKVAITLAGGLAPCLNSAISLERYE 60
Db 1 DVVTWPHLTADIRFCHEPFLNFNFYTLNNKRRKVAITLAGGLAPCLNSAISLERYE 60
Qy 61 IDPSIEIICRYGKGLLGDSYPVTAIEYRKKAGVLOREFGSVIGNSRYKLTNVDCVR 120
Db 61 IDPSIEIICRYGKGLLGDSYPVTAIEYRKKAGVLOREFGSVIGNSRYKLTNVDCVR 120
Qy 121 GLVKEGEDPQKVAADQVYKGVYDIIHTIGDDTNTAAADLAFLARNNYGLTVIGLPTV 180
Db 121 GLVKEGEDPQKVAADQVYKGVYDIIHTIGDDTNTAAADLAFLARNNYGLTVIGLPTV 180
Qy 181 DNDVFPKISLGAMTAAEGGARYFNNVVAENNAENPMLIVHEWGRNGMLTAATQOER 240
Db 181 DNDVFPKISLGAMTAAEGGARYFNNVVAENNAENPMLIVHEWGRNGMLTAATQOER 240
Qy 241 KLIDRAEWLPELGLTRESYEVHAVPEMAIDLEAEAKRLREVMKVCVNI FVSEGAGV 300
Db 241 KLIDRAEWLPELGLTRESYEVHAVPEMAIDLEAEAKRLREVMKVCVNI FVSEGAGV 300
Qy 301 EAIIVAMQAKGOEVRDPAFGHIKLDVNPCKKFGGQFQOMIGAEKTLVOKSGYFAPASAS 360
Db 301 EAIIVAMQAKGOEVRDPAFGHIKLDVNPCKKFGGQFQOMIGAEKTLVOKSGYFAPASAS 360
Qy 361 NVDDMRLLKSCADLAVECAFRESGVIGHDEBNGNVLAIEFPRIKGGKPFNIDTDFNS 420
Db 361 NVDDMRLLKSCADLAVECAFRESGVIGHDEBNGNVLAIEFPRIKGGKPFNIDTDFNS 420
Qy 421 MLSEIGOPKGGKVEVSH 437
Db 421 MLSEIGOPKGGKVEVSH 437
RESULT 4
AAU45865
ID AAU45865 standard; protein; 410 AA.
XX
XX
AC AAU45865;
XX
XX 27-FEB-2002 (first entry)
DT
XX
XX Propionibacterium acnes immunogenic protein #6761.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
XX WO200181581-A2.
PN
XX
XX 01-NOV-2001.
PD
XX
XX 20-APR-2001; 2001WO-US012865.
PF
XX
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI: 2001-616774/71.
DR N-PSDB; AAS59528.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.
XX
XX Example 1, SEQ ID NO 7060, 1065bp, English.
XX
XX Sequences AAU9105-AAU6817 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 410 AA:

Query Match 63.1%; Score 1442; DB 4; Length 410;
Best Local Similarity 68.6%; Pred. No. 3.5e-129;
Matches 271; Conservative 54; Mismatches 70; Indels 0; Gaps 0;

QY 33 KKVALLTGTAGLAPCNSAIGSLIERYTEIDPSIETICRGYKGLLDSDSYVTAENVKK 92
DB 15 KKVALLTGTAGFPAPCLSTIAISGLIQRYTEVAPEVEIIAYKHGEGLLKGFLEVTIVRKN 74
QY 93 AGVLQRFSGSVIGNSRVKLTNVKDCVKRGLVKEGEDPQKVAADQVYKDGVDILHTIGDD 152
DB 75 AEILKRFSGSPIGNSRVKLTNAADLVKRGVLAEGDDPLKVAADRVLADGVDLHTIGDD 134
QY 153 TTTAAADLAFLARNNGYLTIVGLPKTYDNDVFPFKQSLGAWTAABOGARYEMNVAEN 212
DB 135 TTTAAADLAFLAENNYGLTVGLPKTIDNDVPIRQSLGAWTAABOGSRRFQNTVGEHN 194
QY 213 ANPMLTIHEWNGRNCWLTAAOERYKLDRAEMLELGTSTRSYEHAVFVEMAD 272
DB 195 SGRMLIHEWNGRNCWLTAAOERYKLDRAEMLELGTSTRSYEHAVFVEMAD 254
QY 273 LEABAKRLREVNDKVCVIFVSEAGVFAIYAEQAKQOEPRDAFGHIKDDANVPGK 332
DB 255 LEABAKRLKNDKVCVIFVSEAGVFAIYAEQAKQOEPRDAFGHIKDDANVPGK 314
QY 333 FGEQFAQWIGAEKTLVQKSGYPARASNVDMRLIKSCADLAVECAFRESGVIGHDD 392
DB 315 FGKOPADKLGAEKWVQKSGYFSRASAENSEADLELIGRTDLAVOCALGAKGVIGDDE 374
QY 393 NGNVLRATEFPRIKGGKPPNITDWPNSMLSEIQ 427
DB 375 NGDVLTNIAFDRIKGGKPPDITQPFMTLSEIQ 409

RESULT 5
ABM42384
ID ABM42384 standard; protein, 410 AA.

XX AC ABM42384;
XX DT 20-OCT-2003 (first entry)
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #7060.
XX ACNE vulgaris; antibiotic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO200303515-A1.

PD 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

PA (CORI-) CORIXA CORP.

PI Micham JJ, Skeiky YAW, Persing DH, Bhatia A, Malsomeuve JJ;
PT Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglase J;

XX WPI, 2003-381789/36.

DR N-PESDB; ACF64457.

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
or for stimulating an immune response specific for a P. acnes protein.

XX Example 1, SEQ ID NO 7060, 1481bp, English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX encoding a Propionibacterium acnes protein. The invention also relates to
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX immunogenic fragments of P. acnes polypeptides. The invention
XX additionally encompasses expression vectors and host cells comprising a
XX polynucleotide of the invention; antibodies against polypeptides of the
XX invention; fusion proteins comprising a polypeptide of the invention; a
XX method for stimulating an immune response specific for a P. acnes
XX polypeptide and an isolated T cell population comprising T cells prepared
XX via this method; a vaccine composition (comprising P. acnes polypeptides,
XX polynucleotides, antibodies, fusion proteins, T cell populations, or
XX antigen-presenting cells that express the polypeptide); a method and kit
XX for detecting or determining the presence or absence of P. acnes in a
XX patient; and a method for inhibiting the development of P. acnes in a
XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX proteins, T cell populations or antigen-presenting cells that express the
XX polypeptides are useful for diagnosing, preventing or treating acne
XX vulgaris, or for stimulating an immune response specific for a P. acnes
XX protein. The polynucleotides can also be used as probes or primers for
XX nucleic acid hybridization. The vaccine composition is useful for the
XX stimulation of an immune response against P. acnes, or for treating acne,
XX and the kit is useful for performing a diagnostic assay. The present
XX sequence represents a polypeptide predicted to be encoded by an ORF (open
XX reading frame) contained within the P. acnes polynucleotides of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

SO Sequence 410 AA:

Query Match 63.1%; Score 1442; DB 6; Length 410;
Best Local Similarity 68.6%; Pred. No. 3.5e-129;
Matches 271; Conservative 54; Mismatches 70; Indels 0; Gaps 0;

QY 33 KKVALLTGTAGLAPCNSAIGSLIERYTEIDPSIETICRGYKGLLDSDSYVTAENVKK 92
DB 15 KKVALLTGTAGFPAPCLSTIAISGLIQRYTEVAPEVEIIAYKHGEGLLKGFLEVTIVRKN 74
QY 93 AGVLQRFSGSVIGNSRVKLTNVKDCVKRGLVKEGEDPQKVAADQVYKDGVDILHTIGDD 152
DB 75 AEILKRFSGSPIGNSRVKLTNAADLVKRGVLAEGDDPLKVAADRVLADGVDLHTIGDD 134
QY 153 TTTAAADLAFLARNNGYLTIVGLPKTYDNDVFPFKQSLGAWTAABOGARYEMNVAEN 212
DB 135 TTTAAADLAFLAENNYGLTVGLPKTIDNDVPIRQSLGAWTAABOGSRRFQNTVGEHN 194
QY 213 ANPMLTIHEWNGRNCWLTAAOERYKLDRAEMLELGTSTRSYEHAVFVEMAD 272

DB 195 SGSRMLVHEWGRNCGMLTAATPAKYREMLDTQMLPEIGLSKKAWDVHAYVPEAHID 254
 QY 273 LEAEALRLREVMKDVCNIFVSEGAVEAIVAEOMKQGEVPRDAFGHIKLDVAVPGKW 332
 DB 255 LEAEARLRNKMVEGVNITFLSEGAGLDLIEEMEDQGEVRDPFGHKLKNVPGAW 314
 QY 333 FGEQFQOMTGAETLVOKSGYFAPASAVNDMLIKSCADLAVECAFRRSEGVIGHED 392
 DB 315 FGKQFADKGAELKVMQKSGYFERSAASNEADLELIGRCTDLAVDCLAGKTGVIGQDEE 374
 QY 393 NGNVLAIRERRIKGKPFNIDTDWFSMLSEIQ 427
 DB 375 NGDTLTNIAPDRIRKGGKPFDTQTPWFAMLSEIQ 409

RESULT 6
 ADS25315
 ID ADS25315 standard; protein; 184 AA.
 XX
 AC ADS25315;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #14348.
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 EN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI, 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 14348; 122pp; English.
 XX
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 184 AA;
 XX
 XX Query Match 28.9%; Score 659; DB 8; Length 184;
 XX Best Local Similarity 67.9%; Pred. No. 1.4e-54;
 XX Matches 125; Conservative 27; Mismatches 32; Indels 0; Gaps 0;
 XX

QY 33 KKYAIIITAGGLAPCLNSAIGSLIERYTEIDPSIEITCYRGKGLLGSYPYTAVERKK 92
 DB 1 QKYAAMLTAGGLAPCLSSAVGGLIERYSIDIAPEIDIAYRSYGVOGLIGRIEITKDMREK 60
 QY 93 AGVLQRFSGSVIGNSRVKLTNVYDCVKGIAVKEGDPQKVAADOLYKDGVDLHTTGGDD 152
 DB 61 AMLHRYGSGSPIGNSRVKLTNAADCAKGLVKEGDPPLVAERLADGITTHITIGDD 120
 QY 153 TTTAAADLAFLARNYGLTVIGLPKTVNDVFPPIKQSIGAMTTAEQARFPMNVVAENN 212
 DB 121 TTTAAADLAAYLGANCYDITLVGLPKTVNDVFPPIKQSIGAMTTAAVGAAFPNVNEQS 180
 QY 213 ANPR 216
 DB 181 AAPK 184

RESULT 7
 ADN25455
 ID ADN25455 standard; protein; 350 AA.
 XX
 AC ADN25455;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #8108.
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 EN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI, 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a
 XX microbial source, useful for producing plants with improved properties.
 PS Claim 1; SEQ ID NO 8108; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 350 AA;

Query Match 13.4%; Score 306.5; DB 8; Length 350;

Best Local Similarity 26.5%; Pred. No. 2.6e-20;

Matches 104; Conservative 63; Mismatches 145; Indels 81; Gaps 14;

QY 23 FNPYTLNKKKVAITLTAAGLAPCLNSAIGSLIERTEIDPSIEIICRGYKGLIGDS 82
 DB 4 FNEYTOM---RVGVLTGGDCGGLNAVIRAVARRKIK-EYGEYFEGFDGMRGPLEGDT 58
 QY 83 YPVTAEVRKKAGVLOFGGSGVIGNSRVKLTNVKDCVKGGLVKEGEDPQKAAADQVKQGV 142
 DB 59 MFLDIEAVR-GILPR-GGTLIGSSRTNLMKIEGGER-----VQDNNAALGV 103
 QY 143 DILHTITGGDDTTPAADLAFLARNNYGLTVIGLPTVDNDVFPPIKOSIGAWTAAEQGAR 202
 DB 104 DALVAIGGEDTLGVARQL-----HDHGVNVGVPEXTINDNLNATDYTGFDTAIVNATE 157
 QY 203 YEMNVVAENNANPRMLIVHEVMGNRCGWL-----AATQOYRKLLDRAEMLPELGITRE 257
 DB 158 AIDRLHTTAESHRRALVV-EVMGRHAGMATAHAGMAAGA----- 195
 QY 258 SYEVHAVFVPEMAIDLEAEAKRLREVMKVCNIFVSEAGAVEAIVAEQAKQGEVPRD 317
 DB 196 ----NVIILPERPEDIDEVAVYIESFKTNVAPIIVNAGAHPK-----EGQLTLASAEKD 247
 QY 318 AFGHTKLDVAVNPGKWFGEQFAQMIGAETLVOKS---GYPARASNVDDMLIKSCADL 374
 DB 248 SFGHVRLLGSI-----GQRLAEIEIARTGKEARSVVLGHVORGTPSAPFDVLAIRLGIH 301
 QY 375 AVECAFRRESGVIGHDENGNV--LRAIFPRI 405
 DB 302 AITAV-----HDKDFGKVALKGTIVIRV 325

RESULT 8

AB93840

ID ABB93840 standard; protein; 488 AA.

XX ABB93840;

DT 31-MAY-2002 (first entry)

XX

DE Herbicidally active polypeptide SEQ ID NO 3051.

XX Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (PARB) BAYER AG.

PI Tiejien K, Weidner M;

DR WPI: 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 CC organisms.

PS Claim 5; SEQ ID NO 3051; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

XX Sequence 488 AA;

Query Match 12.3%; Score 282; DB 5; Length 488;

Best Local Similarity 26.8%; Pred. No. 9.5e-18;

Matches 112; Conservative 60; Mismatches 158; Indels 88; Gaps 18;

QY 37 ILTAGLAPCLNSAIGSL-----IERTEIDPSIEIICRGYKGLIGDSYPTA 87
 DB 96 IYTCGLCPGLNTYIREVSSLSYWGVRILGIDVSL-----GGYGFYAKNTPIPLNS 149
 QY 88 EVRKKAAGVLOFGGSGVIGNSRVKLTNVKDCVKGGLVKEGEDPQKAAADQVKQGV 147
 DB 150 KV---VNDIHKRGGTLIGTSR-----GSHDTNKL-VDSIQDRIQNYIYI 189
 QY 148 IGGDDTTPAADLAFLARNNYGLTVIGLPTVDNDVFPPIKOSIGAWTAAEQGARF--M 205
 DB 190 IGGDGTGAGAVIPEIRRRRLKVAVGIPTDINDIVIDVSKFSGFDVAVEAQAIVNA 249
 QY 206 NVVAENNANPRMLIVHEVMGNRCGWLTAATQOYRKLLDRAEMLPELGITRESYVAHV 265
 DB 250 HVEAESNENGIGFV--KLMGRYSGIIA-----MYATLASRVDCC 288
 QY 266 VPEMAIDLEAE-----KRLREVMKVCNIFVSEAGAVEAIVAEQAKQGEVPRD 318
 DB 289 IPESPFYIEGEGGLFELERLKHGHNV-----IYLAGAODLMCKMES-----TPMDA 340
 QY 319 FGHITKLDVAVNPGKWFGEQFAQMIGAETLVQ---KSGYPARASNVDDMLIKSCADL 374
 DB 341 SGNKLLKQV--GLWLSQSIKHOFKKNKVMNLKIDPYTMIRAVPSNADSNVCTLLAQ 398
 QY 375 AVECAFRRESGVIGHDE--NGNVIRALIEFPRIKSGKFPNIDT--WFSNMLSEIGOP 428
 DB 399 AVHGAM--AGYGTGSLVNGR--QTYIPFYRITETQNNVYITDRMAR-LISSTNOP 451

RESULT 9

| | | |
|----|--|--|
| AA | AG39480 | |
| ID | AG39480 standard; protein; 485 AA. | |
| XX | | |
| AC | AA39480; | |
| XX | | |
| DT | 18-OCT-2000 (first entry) | |
| DE | Arabidopsis thaliana protein fragment SEQ ID NO: 48654. | |
| XX | | |
| KW | Protein identification; signal transduction pathway; metabolic pathway; | |
| KM | hybridisation assay; genetic mapping; gene expression control; promoter; | |
| XX | termination sequence. | |
| OS | Arabidopsis thaliana. | |
| PN | EPI033405-A2. | |
| XX | | |
| PD | 06-SEP-2000. | |
| FE | 25-FEB-2000; 2000EP-00301439. | |
| XX | | |
| PR | 25-FEB-1999; 99US-0121825P. | |
| PR | 05-MAR-1999; 99US-0123180P. | |
| PR | 09-MAR-1999; 99US-0123548P. | |
| PR | 23-MAR-1999; 99US-0125788P. | |
| PR | 25-MAR-1999; 99US-0126264P. | |
| PR | 29-MAR-1999; 99US-0126785P. | |
| PR | 01-APR-1999; 99US-0127462P. | |
| PR | 06-APR-1999; 99US-0128234P. | |
| PR | 08-APR-1999; 99US-0128714P. | |
| PR | 16-APR-1999; 99US-0129845P. | |
| PR | 19-APR-1999; 99US-0130077P. | |
| PR | 21-APR-1999; 99US-0130449P. | |
| PR | 23-APR-1999; 99US-0130510P. | |
| PR | 23-APR-1999; 99US-0130891P. | |
| PR | 28-APR-1999; 99US-0131449P. | |
| PR | 30-APR-1999; 99US-0132048P. | |
| PR | 04-MAY-1999; 99US-0132407P. | |
| PR | 04-MAY-1999; 99US-0132484P. | |
| PR | 05-MAY-1999; 99US-0132485P. | |
| PR | 06-MAY-1999; 99US-0132486P. | |
| PR | 06-MAY-1999; 99US-0132487P. | |
| PR | 07-MAY-1999; 99US-0132863P. | |
| PR | 11-MAY-1999; 99US-0134256P. | |
| PR | 14-MAY-1999; 99US-0134218P. | |
| PR | 14-MAY-1999; 99US-0134219P. | |
| PR | 14-MAY-1999; 99US-0134221P. | |
| PR | 14-MAY-1999; 99US-0134370P. | |
| PR | 18-MAY-1999; 99US-0134768P. | |
| PR | 19-MAY-1999; 99US-0134941P. | |
| PR | 20-MAY-1999; 99US-0135124P. | |
| PR | 21-MAY-1999; 99US-0135353P. | |
| PR | 24-MAY-1999; 99US-0135629P. | |
| PR | 25-MAY-1999; 99US-0136021P. | |
| PR | 27-MAY-1999; 99US-0136392P. | |
| PR | 28-MAY-1999; 99US-0136782P. | |
| PR | 01-JUN-1999; 99US-0137232P. | |
| PR | 03-JUN-1999; 99US-0137528P. | |
| PR | 04-JUN-1999; 99US-0137502P. | |
| PR | 07-JUN-1999; 99US-0137724P. | |
| PR | 08-JUN-1999; 99US-0138094P. | |
| PR | 10-JUN-1999; 99US-0138540P. | |
| PR | 10-JUN-1999; 99US-0138847P. | |
| PR | 14-JUN-1999; 99US-0139119P. | |
| PR | 16-JUN-1999; 99US-0139452P. | |
| PR | 16-JUN-1999; 99US-0139453P. | |
| PR | 17-JUN-1999; 99US-0139492P. | |
| PR | 18-JUN-1999; 99US-0139454P. | |
| PR | 18-JUN-1999; 99US-0139455P. | |
| PR | 18-JUN-1999; 99US-0139456P. | |
| PR | 18-JUN-1999; 99US-0139457P. | |
| PR | 18-JUN-1999; 99US-0139458P. | |
| PR | 18-JUN-1999; 99US-0139459P. | |

| | | |
|----|-------------|----------------|
| PR | 18-JUN-1999 | 99US-01394602 |
| PR | 18-JUN-1999 | 99US-0139461P |
| PR | 18-JUN-1999 | 99US-0139462P |
| PR | 18-JUN-1999 | 99US-0139463P |
| PR | 18-JUN-1999 | 99US-0139463P |
| PR | 18-JUN-1999 | 99US-0139475P |
| PR | 18-JUN-1999 | 99US-0139476P |
| PR | 21-JUN-1999 | 99US-0139487P |
| PR | 22-JUN-1999 | 99US-0139489P |
| PR | 23-JUN-1999 | 99US-0140035P |
| PR | 23-JUN-1999 | 99US-0140034P |
| PR | 24-JUN-1999 | 99US-0140695P |
| PR | 28-JUN-1999 | 99US-0140823P |
| PR | 29-JUN-1999 | 99US-0140991P |
| PR | 30-JUN-1999 | 99US-0141827P |
| PR | 01-JUL-1999 | 99US-0141842P |
| PR | 01-JUL-1999 | 99US-0142154P |
| PR | 01-JUL-1999 | 99US-0142055P |
| PR | 06-JUL-1999 | 99US-0142290P |
| PR | 08-JUL-1999 | 99US-0142803P |
| PR | 09-JUL-1999 | 99US-0142920P |
| PR | 12-JUL-1999 | 99US-0142977P |
| PR | 13-JUL-1999 | 99US-0143542P |
| PR | 14-JUL-1999 | 99US-0143624P |
| PR | 15-JUL-1999 | 99US-0144005P |
| PR | 16-JUL-1999 | 99US-0144085P |
| PR | 16-JUL-1999 | 99US-0144308P |
| PR | 19-JUL-1999 | 99US-0144335P |
| PR | 19-JUL-1999 | 99US-0144332P |
| PR | 19-JUL-1999 | 99US-0144331P |
| PR | 19-JUL-1999 | 99US-0144332P |
| PR | 19-JUL-1999 | 99US-0144333P |
| PR | 19-JUL-1999 | 99US-0144334P |
| PR | 19-JUL-1999 | 99US-0144335P |
| PR | 20-JUL-1999 | 99US-0144352P |
| PR | 20-JUL-1999 | 99US-0144632P |
| PR | 20-JUL-1999 | 99US-0144684P |
| PR | 21-JUL-1999 | 99US-0144684P |
| PR | 21-JUL-1999 | 99US-0145086P |
| PR | 21-JUL-1999 | 99US-0145088P |
| PR | 22-JUL-1999 | 99US-0145085P |
| PR | 22-JUL-1999 | 99US-0145087P |
| PR | 22-JUL-1999 | 99US-0145089P |
| PR | 22-JUL-1999 | 99US-0145192P |
| PR | 23-JUL-1999 | 99US-0145145P |
| PR | 23-JUL-1999 | 99US-0145218P |
| PR | 23-JUL-1999 | 99US-0145224P |
| PR | 25-JUL-1999 | 99US-0145276P |
| PR | 27-JUL-1999 | 99US-0145913P |
| PR | 27-JUL-1999 | 99US-0145918P |
| PR | 27-JUL-1999 | 99US-0145919P |
| PR | 28-JUL-1999 | 99US-0145951P |
| PR | 02-AUG-1999 | 99US-0146386P |
| PR | 02-AUG-1999 | 99US-0146388P |
| PR | 02-AUG-1999 | 99US-0146389P |
| PR | 03-AUG-1999 | 99US-0147038P |
| PR | 04-AUG-1999 | 99US-0147204P |
| PR | 04-AUG-1999 | 99US-0147302P |
| PR | 05-AUG-1999 | 99US-0147329P |
| PR | 05-AUG-1999 | 99US-0147260P |
| PR | 06-AUG-1999 | 99US-0147303P |
| PR | 06-AUG-1999 | 99US-0147436P |
| PR | 09-AUG-1999 | 99US-0147453P |
| PR | 09-AUG-1999 | 99US-0147935P |
| PR | 10-AUG-1999 | 99US-0148171P |
| PR | 11-AUG-1999 | 99US-0148319P |
| PR | 12-AUG-1999 | 99US-0148311P |
| PR | 13-AUG-1999 | 99US-0148355P |
| PR | 13-AUG-1999 | 99US-0148684P |
| PR | 16-AUG-1999 | 99US-0149368P |
| PR | 17-AUG-1999 | 99US-01494715P |
| PR | 18-AUG-1999 | 99US-0149426P |
| PR | 20-AUG-1999 | 99US-0149722P |
| PR | 20-AUG-1999 | 99US-0149723P |
| PR | 20-AUG-1999 | 99US-0149929P |

```

PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158332P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

Query Match 12.3%; Score 281.5; DB 3; Length 485;
 Best Local Similarity 26.7%; Pred. No. 1.1e-17;
 Matches 109; Conservative 61; Mismatches 166; Indels 73; Gaps 16;

```

QY 37 ILTAGGLACGSAAGSLIERTEIDPSLEIICVGGYGLLGDSPYTAVRKKAGVL 96
DB 96 ITCGGLCGALNTVREVVSSLYMGVRRILGIDGGYGYAKNTIPINSKV---VNDI 152
QY 97 QRFSGSVIGNSRKLTVNVDQVRGLVKEGEPDQKVAADQLVKGVDLHTIGDPTNTA 156
DB 153 HRRGGTITTSR-----GSHDTNKKI-VDSIDRGINQVYIIGSGDTORG 195
QY 157 AADLAAPFLARNYGLTVGLPTVDNDFPIKOSIGCAWTAAGARF--NMVVAENNAN 214
DB 196 ASVIFEEIRRRRLKLVAVGIPKTIIDNDIPIDKSFQFDPAVEBAQAIYAAHVEAESNEN 255
QY 215 PRLMLVHEVMGNCGMLTAATQAEYKLLDRAEMLEPGLTNESEYVAHVPEBAIDLE 274

```

```

DB 256 GIGFV--KMGRTYSCYIA-----MTATLASRDVDCCLIPSEPFYIE 294
QY 275 AEA-----KRLREVMKVCWNIFVSEGAVEAIVAEQAKGQEVPRDAFGH1KLDAY 327
DB 295 GEGGLFEFIERLRKDHGMV-----IVLAGAGQDLMCKSMES-----TPMDASGNKLLKOV 346
QY 328 NPGKWFGEQFAQMIGAEKTLVQ---KSGYFARASAVDDKRLKSCADLAVECAFRRE 383
DB 347 --GLWLSQSIDHFKKKNKVNKLKYIDPTYMIKAVPSNASDQVYCTLLAQSAVHGAM--- 401
QY 364 SGVIGHDD--NGNVLRAIEFPRIKGRPFNIDT--MFSNWLSEIGP 428
DB 402 AGYTGYSGLVNGR-QTYIIPFRITETQNNVITDRMAR-LLSSTNDP 448

```

RESULT 10
 ID ABU19327 standard; protein; 448 AA.
 AC ABU19327;
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #4854.
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS Borrelia burgdorferi.
 PN WO200277183-A2.
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002WO-US009107.
 PR 21-MAR-2001; 2001US-00945242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.

XX (ELT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA23197.
 DR
 XX
 PT New antisense nucleic acid, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 47251; 1766bp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 623 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX Oryza sativa.
OS Key Location/Qualifiers
FH Misc-difference 130 /note= "encoded by codon AAT"
FT
XX
XX MO9505457-A1.
XX
XX 23-FEB-1995.
XX
XX 16-AUG-1994; 94WO-JP001352.
XX PF
XX 19-AUG-1993; 93JP-00226454.
XX
XX (NISR) JAPAN TOBACCO INC.
XX
XX Hiyoeshi T, Mine T, Kasaoka K, Tyson HR, Page MJA;
XX
XX WPI, 1995-098757/13.
XX DR
XX N-PSDB; AAQ85984.
XX
XX DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
PT origin, for prodn. of transformant plant cells with altered sugar
PT content.
XX
XX Claim 7, Page 46-49; 79pp; Japanese.
XX
XX The amino acid sequence of the novel ATP-dependent fructose-6-phosphate 1
CC -phosphotransferase enzyme (EC 2.7.1.11; PFK) encoded by the Oryza sativa
CC (rice) gene, pPFK-OS, is given in the specification. Plants transformed
CC with genes encoding PFK (see AAQ85982-86) can express the enzyme. The
CC transformed plants can produce varieties that have altered sugar content
CC on storage at low temperatures. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
XX Sequence 469 AA;
SQ
Query Match 11.8%; Score 269.5; DB 2; Length 469;
Best Local Similarity 26.8%; Pred. No. 1.4e-16;
Matches 109; Conservative 57; Mismatches 172; Indels 69; Gaps 16;
QY 37 ILTAGLAPCLNSALGSLIERTEIDPSIEITCYRGYGLLGDSPYTAVERKAGVL 96
DB 74 IVTGGLCPGLNTVIREIVCGLDVYGVSRVLGIGGGRGYFACTIDLSP---KSVNDI 130
QY 97 ORFGGSVIGNSRVKLTNVKDCVKGGLVKEGEDPQKVAADOLVKQGVDLHTIGDPTNTA 156
DB 131 HRRGGTIVGTSR-----GHDYTKI-VDSDRGINOVYIIGSGDGTQRG 173
QY 157 AADLAAPLARNNYGLTVIGLPKTVNDVFPRIKOSLGAWTAAEQARYFMNVVAENNANPR 216
DB 174 AGVIEEIRRRGLKAVAVAGIPKTIINDIPIDRSFGPTVAEBQRAINAAHVEAGSAEN 233
QY 217 MLIYHEVNGRNGCWLTAATAQERYKLDRAEWLPGLGLTRESYEVHAFVPEMAIDLEAE 276
DB 234 GIGLVKLMGRHSGFLA---HYATLASR-----DVDCCIIPESSPEFLRGE 274
QY 277 A-----KRLAEVNDKDCVNIIFYSEGAGVEAIVAEQAQGEVPRDAFGH-IKLDVAN 328
DB 275 GGLFYLELGRLEKNGHNV---IYVABEGGQKLIINETKESMG---KDAISGSLIDLV-- 324
QY 329 PCKWGEOPAQWIGAEKTLVQ---KSGYFARASASNVDDMLIKSCADLAVECAFRRBS 384
DB 325 -GLWISQRIKHEPKIKTITINLKTYIDPTVMIRAIPSNADNVYCTLLAHVVHAGMAGYT 383
QY 385 G-VIGHDENGVLAIEPRPI--KGGKPFNIDTOWFNSMLSEIQOP 428
DB 384 GPTVG--QVNGRHC-YIPFYRIITEKONKVSINDRWAR-LIESTTOP 426

ID AAR71583 standard; protein; 426 AA.
XX
XX AAR71583;
AC
XX 25-MAR-2003 (revised)
DT 12-OCT-1995 (first entry)
XX
XX Raphanus sativus fructose-6-phosphate 1-phosphotransferase.
DE
XX
XX ATP-dependent fructose-6-phosphate 1-phosphotransferase enzyme; plant;
KM potato; Solanum tuberosum; rice; Oryza sativa; maize; Zea mays; radish;
KM Raphanus sativus; Flaveria brownii; primer: expression vector;
KM Agrobacterium tumefaciens; sugar; storage; temperature.
XX
XX Raphanus sativus.
OS
XX MO9505457-A1.
XX
XX 23-FEB-1995.
XX
XX 16-AUG-1994; 94WO-JP001352.
XX PF
XX 19-AUG-1993; 93JP-00226454.
XX PR
XX (NISR) JAPAN TOBACCO INC.
XX
XX Hiyoeshi T, Mine T, Kasaoka K, Tyson HR, Page MJA;
XX
XX WPI, 1995-098757/13.
XX DR
XX N-PSDB; AAQ85986.
XX
XX DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
PT origin, for prodn. of transformant plant cells with altered sugar
PT content.
XX
XX Claim 11, Page 58-61; 79pp; Japanese.
XX
XX The amino acid sequence of the novel ATP-dependent fructose-6-phosphate 1
CC -phosphotransferase enzyme (EC 2.7.1.11; PFK) encoded by the Raphanus
CC sativus (radish) gene, pPFK-RS1. Plants transformed with genes encoding
CC PFK (see AAQ85982-86) can express the enzyme. The transformed plants can
CC produce varieties that have altered sugar content on storage at low
CC temperatures. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 426 AA;
SQ
Query Match 11.5%; Score 263; DB 2; Length 426;
Best Local Similarity 25.8%; Pred. No. 5.2e-16;
Matches 108; Conservative 65; Mismatches 173; Indels 72; Gaps 17;
QY 37 ILTAGLAPCLNSALGSLIERTEIDPSIEITCYRGYGLLGDSPYTAVERKAGVL 96
DB 50 IVTGGLCPGLNTVIREIVCGLSVYGVKILIGEGRGYFARNTIDL---KTVNDI 106
QY 97 ORFGGSVIGNSRVKLTNVKDCVKGGLVKEGEDPQKVAADOLVKQGVDLHTIGDPTNTA 156
DB 107 HRRGGTIVGTSR-----GHDYTKI-VDSDRGINOVYIIGSGSGKX 149
QY 157 AADLAAPLARNNYGLTVIGLPKTVNDVFPRIKOSLGAWTAAEQARYFMNVVAENNANPR 216
DB 150 AAVIEEIRRRGLKAVAVAGIPKTIINDIPIDRSFGPTVAEBQRAINAAHVEATSTEN 209
QY 217 MLIYHEVNGRNGCWLTAATAQERYKLDRAEWLPGLGLTRESYEVHAFVPEMAIDLEAE 276
DB 210 GIGLVKLMGRHSGFLA---HYATLASR-----DVDCCIIPESSPEFLRGE 250
QY 277 -----AKRLAEVNDKDCVNIIFYSEGAGVEAIVAEQAQGEVPRDAFGH-IKLDVAN 329
DB 251 GGLFYLELGRLEKNGHNV---IYVABEGGQKLIINETKESMG---KDAISGSLIDLV-- 302
QY 330 GKMFGGEOPAQWIGAEKTLVQ---SGYFARASASNVDDMLIKSCADLAVECAFRRBSGV 386
DB 303 GLWISQRIKHEPKIKTITINLKTYIDPTVMIRAIPSNADNVYCTLLAHVVHAGMAGYT 355

Qy 387 I-GHEDNGNVLA-----IEFPRI--KGKPFNIIDTFWNSMLSEIGQPKGKVEVSH 437
 Db 356 MAGYNGFTVGLVGRHTYIPFYRITEKONKVVITDRMMAR-LLSTNQSPFMGDDHH 412

RESULT 14

ID AAR71579 standard; protein; 485 AA.

AC AAR71579;

DT 25-MAR-2003 (revised)
 DT 11-OCT-1995 (first entry)

DE Solanum tuberosum fructose-6-phosphate 1-phosphotransferase.

KM ATP-dependent fructose-6-phosphate 1-phosphotransferase enzyme; plant;
 KM potato; Solanum tuberosum; rice; Oryza sativa; maize; Zea mays; radish;
 KM Raphanus sativus; Flaveria browningii; primer; expression vector;
 KM Agrobacterium tumefaciens; sugar; storage; temperature.

OS Solanum tuberosum.

PN WO9505457-A1.

PD 23-FEB-1995.

PF 16-AUG-1994; 94NO-JP001352.

PR 19-AUG-1993; 93JP-00226454.

PA (NISR) JAPAN TOBACCO INC.

PI Hiyeohi T, Mine T, Kasaka K, Tyson HR, Page MD.

DR WPI; 1995-098757/13.

DR N-PSDB; AAQ85982.

PT DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant
 origin, for prodn. of transformant plant cells with altered sugar
 content.

PS Claim 4; Page 33-37; 79pp; Japanese.

CC The amino acid sequence of the novel ATP-dependent fructose-6-phosphate 1
 CC -phosphotransferase enzyme (EC 2.7.1.11; PFK) encoded by the potato
 CC (Solanum tuberosum) gene PFK32. The DNA sequence was isolated from a
 CC potato cDNA library using primers AAQ85987-91. The sequence was used to
 CC construct the plant expression vector PPK(35S) which was transformed
 CC into potato plants (Bintje variety) via Agrobacterium tumefaciens. The
 CC transformed plants can express the enzyme and produce plant varieties
 CC that have altered sugar content on storage at low temperatures. (Updated
 CC on 25-MAR-2003 to correct PN field.)

CC Sequence 485 AA;

Query Match 11.5%; Score 262; DB 2; Length 485;

Best Local Similarity 25.7%; Pred. No. 7.8e-16; Indels 92; Gaps 17;

Matches 113; Conservative 56; Mismatches 179; Indels 92; Gaps 17;

Qy 7 YHLDADIRCFHMFLLNFNFYTLMMKPKVAIILTAGGLACPLNSAIGSLIERYTEIDPSIE 66

Db 86 YRSSDDVRAC-----IVTCGLCPRLNIVIEIHSLDYMGVAK 125

Qy 67 IICRYGGYGLLIGDSYPTAENRKAGVLRFGSSVIGNSRVKLTNVKDCVRLGKEG 126

Db 126 VFGIDGGYGFYSKNILNLP--KTVNDIHKRGCTILGSSR-----CG 166

Qy 127 EDPQKVAADQVLVDGDIILHTTGGDDNTAAADLAFIARNNGLVIGLPTVDNDVFP 186

Db 167 HDTTKI-VDSIQDREINOVYIIGDGTQKGAAYIEIRRGKLVIVAGIPTINDIPV 225

Qy 187 IKSLGAMTAAQAGAFYFNWVAENNANPRLMIVHEMGRNCGLTAATAQERYKLIDRA 246
 Db 226 IDKSFQDFAVEBAQAIINAAHVEASANGICVWLMGRVSGFIA----- 271
 Qy 247 EMLPELGLTRESVEYNAVVEPAIDLEAA-----KRLREVMDQCAIIFVEGAG 299
 Db 272 -----MYATLASADVDLCLIPESPFYLEDGGLFEYIERLKENGMV-----IVIAEGAG 322
 Qy 300 VEALVAE-MQAKQOEPRDAFGHIKLDVANPCKMFGEOFAOMIGAEKTLVQ-----KSG 352
 Db 323 QELIAEENMAAKNEQ---DASGNKLIQDY--GLMISQKIRDNH-ARKTGMPTLKYIDPF 376
 Qy 353 YFAPASASVDDMRILIKSCADLAVECAFRRSGVIGHDE--NGNVLRAIEFPRIKGGKP 410
 Db 377 YMIRAVPSNASDN---VYCTLLAQSCVHGAMAGYFTSGLVNGR-QTYIPFNRITEKON 432
 Qy 411 FNIDTD--WFNSMLSEIGP 428
 Db 433 MVVITDRMMAR-LLSTNQP 451

RESULT 15

ID ADC07918 standard; protein; 496 AA.

AC ADC07918;

DT 18-DEC-2003 (first entry)

DE Rice protein sequence Seg ID184 related to grain filling.

KM plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 KM carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 KM tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
 KM wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
 KM gene; ds; plant.

OS Oryza sativa.

PN MO2003000905-A2.

PR 22-JUN-2001; 2001US-0300112P.

PR 26-SEP-2001; 2001US-0325277P.

PR 20-DEC-2001; 2001US-0342327P.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;

PI Glazebrook J, Katagiri F, Kreps J, Provart N, Riecke D;

DR N-PSDB; ADC07917.

DR WPI; 2003-229341/22.

PT New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.

PS Claim 1; SEQ ID NO 184; 130pp; English.

CC This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
 CC sugarbeet, wheat, and rice. The invention may be useful for the
 CC improvement of protein, oil, starch, fibre and moisture content of the
 CC cereal grains. In addition, carbohydrate levels may be modified to a more

CC desirable level using the present invention. The present sequence is the
CC amino acid sequence of a rice protein of the invention. Note: the
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/publishedpct_sequences.

XX
SQ Sequence 496 AA;

Query Match 11.4%; Score 261.5; DB 7; Length 496;
Best Local Similarity 26.2%; Pred. No. 9e-16;
Matches 107; Conservative 54; Mismatches 180; Indels 67; Gaps 13;

```
OY 36 ALTTAGGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSPVPTAEVRKXGV 95
DB 95 AIVTCGGICPGELNIVRELVCGLYDMYGVTSVVGIEGGYKGYSRVYALTP--KSVND 151
OY 96 LORFGGSVIGNSRVKLTNVKDCVKRGLYKGEDPQKVAADQLVKDGVDLHTIGGDDTNT 155
DB 152 IHKRGSTVLGTSR-----CGHDTGKI-VDSIKDRGINQVYIIIGGDTQK 194
OY 156 AAADLAFLARNNGYLTIGLPKTYDNDVFPIKQSLGAWTAAEGARFPMNVVAENNANP 215
DB 195 GASVIYEVEVRRRGLKCSVVGVPKTDINDIAVIDKSFSGFDTAVEAQRAINAAVEAESAE 254
OY 216 RMLIVHEVWGRNCGWLTATAQRYKLDRAEMLPBLGLTRESYEHAVFVPEMAIDLEA 275
DB 255 NGIGVVKLMGRNSGFIA-----MVAATLASRDVDCCLIPESPPYLEG 295
OY 276 EA-----KRLREVMWDKVCVNIIFVSEAGYEAIVAEWQAKGOEVPKDAFGHIKLDVAVN 328
DB 296 KGGLEEFIEKRLKDNQHNV-----IYVABEGAGDILAKSNPFVDTQ--DASGNKLL--LD 346
OY 329 PKGMFGEOPAQMTGAETL-----VQKSGYFAPASASNTVDMRLIKSCADLAVECAFRRR 383
DB 347 VGLMLSQKIKDHPKKRNFPITLKYIDPYMIRAVRSNASDN---VYCTLLAHSAHLGAM 403
OY 384 SGVIGHDEDNQNLRA-IEFPRI--KGGKPFNIDTDMFMSMLSEIGOP 428
DB 404 AGYTGFTVAPVNGRHAYIPFYRITEKONKRVITTRMMAR-VLCSTNOP 450
```

Search completed: February 24, 2005, 14:19:27
Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 14:13:35 ; Search time 44 Seconds
(without alignments)
741.401 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTPYHLLTDIRCFHWF.....FNSMLEIGQPKGKGVESH 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patente AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---|
| 1 | 2284 | 100.0 | 437 | 4 | US-09-934-901-16 Sequence 16, Appl |
| 2 | 2284 | 100.0 | 437 | 4 | US-09-934-868-6 Sequence 6, Appl |
| 3 | 2284 | 100.0 | 437 | 4 | US-10-321-210-16 Sequence 16, Appl |
| 4 | 2284 | 100.0 | 437 | 4 | US-10-320-874-16 Sequence 16, Appl |
| 5 | 270 | 11.8 | 483 | 2 | US-08-416-870C-4 Sequence 4, Appl |
| 6 | 264.5 | 11.6 | 469 | 2 | US-08-416-870C-6 Sequence 6, Appl |
| 7 | 262 | 11.5 | 485 | 2 | US-08-416-870C-2 Sequence 2, Appl |
| 8 | 261 | 11.4 | 445 | 4 | US-09-902-540-16136 Sequence 16136, A |
| 9 | 260 | 11.4 | 426 | 2 | US-08-416-870C-10 Sequence 10, Appl |
| 10 | 227 | 9.9 | 338 | 4 | US-09-489-039A-12258 Sequence 12258, A |
| 11 | 226.5 | 9.9 | 345 | 3 | US-09-134-001C-31774 Sequence 31774, A |
| 12 | 220 | 9.6 | 330 | 4 | US-09-543-681A-8337 Sequence 8337, Ap |
| 13 | 220 | 9.6 | 335 | 4 | US-09-583-110-5099 Sequence 5099, Ap |
| 14 | 218.5 | 9.6 | 321 | 4 | US-09-134-000C-5332 Sequence 5332, Ap |
| 15 | 215.5 | 9.4 | 323 | 4 | US-09-107-532A-6470 Sequence 6470, Ap |
| 16 | 211.5 | 9.3 | 522 | 2 | US-08-416-870C-8 Sequence 8, Appl |
| 17 | 210 | 9.2 | 389 | 4 | US-09-849-016-9488 Sequence 9488, Ap |
| 18 | 206.5 | 9.0 | 781 | 1 | US-08-280-690-2 Sequence 2, Appl |
| 19 | 204.5 | 9.0 | 784 | 4 | US-09-849-016-6251 Sequence 6251, Ap |
| 20 | 201.5 | 8.8 | 798 | 4 | US-09-849-016-11096 Sequence 11096, A |
| 21 | 201 | 8.8 | 343 | 4 | US-09-887-054-2 Sequence 2, Appl |
| 22 | 182 | 8.0 | 307 | 4 | US-09-710-279-2122 Sequence 2122, Ap |
| 23 | 181 | 7.9 | 550 | 4 | US-09-198-452A-225 Sequence 225, Ap |
| 24 | 181 | 7.9 | 556 | 4 | US-09-438-185A-210 Sequence 210, Ap |
| 25 | 179 | 7.8 | 556 | 4 | US-09-438-185A-162 Sequence 162, Ap |
| 26 | 171.5 | 7.5 | 544 | 4 | US-09-198-452A-180 Sequence 180, Ap |
| 27 | 148 | 6.5 | 696 | 4 | US-09-949-016-8015 Sequence 8015, Ap |

| | | | | | |
|----|-------|-----|------|---|--|
| 28 | 120 | 5.3 | 169 | 4 | US-09-107-433-2681 Sequence 2681, Ap |
| 29 | 113.5 | 5.0 | 845 | 4 | US-09-198-452A-458 Sequence 458, App |
| 30 | 113.5 | 5.0 | 847 | 4 | US-09-438-185A-439 Sequence 439, App |
| 31 | 107 | 4.7 | 810 | 4 | US-09-583-110-4352 Sequence 4352, Ap |
| 32 | 107 | 4.7 | 816 | 4 | US-09-107-433-4504 Sequence 4504, Ap |
| 33 | 106.5 | 4.7 | 543 | 4 | US-09-248-796A-17070 Sequence 17070, Ap |
| 34 | 101.5 | 4.4 | 1427 | 4 | US-09-538-092-1044 Sequence 1044, Ap |
| 35 | 100.5 | 4.4 | 341 | 4 | US-09-248-796A-17051 Sequence 17051, A |
| 36 | 99.5 | 4.4 | 742 | 4 | US-09-489-039A-12100 Sequence 12100, A |
| 37 | 99.5 | 4.4 | 1038 | 3 | US-09-541-782-4 Sequence 4, Appl |
| 38 | 99.5 | 4.4 | 1038 | 4 | US-09-723-820-4 Sequence 4, Appl |
| 39 | 99.5 | 4.4 | 1038 | 4 | US-10-270-085-4 Sequence 4, Appl |
| 40 | 97.5 | 4.3 | 413 | 4 | US-09-603-208A-8 Sequence 8, Appl |
| 41 | 97.5 | 4.3 | 538 | 4 | US-09-603-208A-6 Sequence 6, Appl |
| 42 | 97.5 | 4.3 | 724 | 4 | US-09-489-039A-12100 Sequence 12100, A |
| 43 | 96.5 | 4.2 | 458 | 4 | US-09-489-039A-9184 Sequence 9184, Ap |
| 44 | 96 | 4.2 | 863 | 4 | US-09-328-352-6730 Sequence 6730, Ap |
| 45 | 95.5 | 4.2 | 813 | 4 | US-09-252-991A-30252 Sequence 30252, A |

ALIGNMENTS

RESULT 1
US-09-934-901-16
Sequence 16, Application US/09934901
Patent No. 655353
GENERAL INFORMATION:
APPLICANT: Kotfas, Matheos
APPLICANT: Odum, J. Martin
APPLICANT: No. 65535310n, Kelley C.
APPLICANT: Ye, Rick
TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: C1619 US NA
CURRENT APPLICATION NUMBER: US/09/934,901
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,906
PRIOR FILING DATE: September 1, 2000
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Microsoft Office 97
SEQ ID NO 16
LENGTH: 437
TYPE: PR
ORGANISM: METHYLOMONAS SP.
US-09-934-901-16

| | | | | |
|-----------------------|----------------|--|--|------------|
| Query Match | 100.0% | Score 2284 | DB 4 | Length 437 |
| Best Local Similarity | 100.0% | Pred. No. 1.6e-215 | | |
| Matches 437 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 1 | DVVTTPYHLLTDIRCFHWF | FLNENFYTLNKKRKAAILTAGGLACLSAIGSLIERYTE | 60 |
| DB | 1 | DVVTTPYHLLTDIRCFHWF | FLNENFYTLNKKRKAAILTAGGLACLSAIGSLIERYTE | 60 |
| QY | 61 | IDPSLEIICRYRGYGLLLGDSYPTAEVRKAGVLOFSGSVIGNSRYKLTVNDCVGR | 120 | |
| DB | 61 | IDPSLEIICRYRGYGLLLGDSYPTAEVRKAGVLOFSGSVIGNSRYKLTVNDCVGR | 120 | |
| QY | 121 | GLVKEGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV | 180 | |
| DB | 121 | GLVKEGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV | 180 | |
| QY | 181 | DNDVPRIPQISGAMTPAAAGAGAYFNNVVAENNANRMLIVHEVMGRNCGLTAAATROEYR | 240 | |
| DB | 181 | DNDVPRIPQISGAMTPAAAGAGAYFNNVVAENNANRMLIVHEVMGRNCGLTAAATROEYR | 240 | |
| QY | 241 | KLLDRAEMLPELGLTRESYEVHAFVPEMAIDLEAEAKRLREVMKVCVNFVSEGAGV | 300 | |
| DB | 241 | KLLDRAEMLPELGLTRESYEVHAFVPEMAIDLEAEAKRLREVMKVCVNFVSEGAGV | 300 | |
| QY | 301 | EAIIVAEQKQGEVPRDAFGHIKLDAVNPCKKFGQFOPMIGAEKTLVQKSGYFAPASAS | 360 | |
| DB | 301 | EAIIVAEQKQGEVPRDAFGHIKLDAVNPCKKFGQFOPMIGAEKTLVQKSGYFAPASAS | 360 | |

```
QY 361 NVDDMRLLKSCADLAVECAFRESGVIGHDENGVLAIEFPRIKGGKPFNIDTWENS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVIGHDENGVLAIEFPRIKGGKPFNIDTWENS 420
QY 421 MLSEIGOPKGGKVEVESH 437
DB 421 MLSEIGOPKGGKVEVESH 437

RESULT 2
US-09-934-868-6
; Sequence 6, Application US/09934868
; Patent No. 6689601
; GENERAL INFORMATION:
; APPLICANT: Koffas, Matheos
; APPLICANT: Odem, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
; US-09-934-868-6

Query Match 100.0%; Score 2284; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-215;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVVTMPYHLTADIRFCHEFLNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVVTMPYHLTADIRFCHEFLNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICRGYKGLLGDSYPTVAEVRKKAAGVLOFGSGVIGNSRVKLTNVDCVVR 120
DB 61 IDPSIEIICRGYKGLLGDSYPTVAEVRKKAAGVLOFGSGVIGNSRVKLTNVDCVVR 120
QY 121 GLVKEGEDPQKVAADQVLDVGDILHTTIGDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKEGEDPQKVAADQVLDVGDILHTTIGDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGWLTAAATAOEYR 240
DB 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGWLTAAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRSEYEVHAFVPEMAIDLEAARLREVMKVDVCNIFVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRSEYEVHAFVPEMAIDLEAARLREVMKVDVCNIFVSEGAGV 300
QY 301 EAIIVAMQAKQOEVRDAFGHIKLDVNPCKMFGSEQFAQMIGAEXTLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQOEVRDAFGHIKLDVNPCKMFGSEQFAQMIGAEXTLVOKSGYFARASAS 360
QY 361 NVDDMRLLKSCADLAVECAFRESGVIGHDENGVLAIEFPRIKGGKPFNIDTWENS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVIGHDENGVLAIEFPRIKGGKPFNIDTWENS 420
QY 421 MLSEIGOPKGGKVEVESH 437
DB 421 MLSEIGOPKGGKVEVESH 437

RESULT 3
US-10-321-210-16
; Sequence 16, Application US/10321210
; Patent No. 6767744
```

```
; GENERAL INFORMATION:
; APPLICANT: Koffas, Matheos
; APPLICANT: Odem, J. Martin
; APPLICANT: No. 6767744ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/321,210
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
; US-10-321-210-16

Query Match 100.0%; Score 2284; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-215;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVVTMPYHLTADIRFCHEFLNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVVTMPYHLTADIRFCHEFLNPFYTLNKPCKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICRGYKGLLGDSYPTVAEVRKKAAGVLOFGSGVIGNSRVKLTNVDCVVR 120
DB 61 IDPSIEIICRGYKGLLGDSYPTVAEVRKKAAGVLOFGSGVIGNSRVKLTNVDCVVR 120
QY 121 GLVKEGEDPQKVAADQVLDVGDILHTTIGDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKEGEDPQKVAADQVLDVGDILHTTIGDDTNTAAADLAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGWLTAAATAOEYR 240
DB 181 DNDVFPRIKOSLGAMTAAEGARFEMNVVAENNANPRMLIVHEVMGRNGWLTAAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRSEYEVHAFVPEMAIDLEAARLREVMKVDVCNIFVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRSEYEVHAFVPEMAIDLEAARLREVMKVDVCNIFVSEGAGV 300
QY 301 EAIIVAMQAKQOEVRDAFGHIKLDVNPCKMFGSEQFAQMIGAEXTLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQOEVRDAFGHIKLDVNPCKMFGSEQFAQMIGAEXTLVOKSGYFARASAS 360
QY 361 NVDDMRLLKSCADLAVECAFRESGVIGHDENGVLAIEFPRIKGGKPFNIDTWENS 420
DB 361 NVDDMRLLKSCADLAVECAFRESGVIGHDENGVLAIEFPRIKGGKPFNIDTWENS 420
QY 421 MLSEIGOPKGGKVEVESH 437
DB 421 MLSEIGOPKGGKVEVESH 437

RESULT 4
US-10-320-874-16
; Sequence 16, Application US/10320874
; Patent No. 6773905
; GENERAL INFORMATION:
; APPLICANT: Koffas, Matheos
; APPLICANT: Odem, J. Martin
; APPLICANT: No. 6773905ton, Kelley C.
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/320,874
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
```

PRIOR APPLICATION NUMBER: 60/229,906
 PRIOR FILING DATE: September 1, 2000
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: Microsoft Office 97
 SEQ ID NO: 16
 LENGTH: 437
 TYPE: PRT
 ORGANISM: METHYLOMONAS SP.
 US-10-320-874-16

Query Match 100.0%; Score 2284; DB 4; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1,66-215;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVATWPHYLTADIRFCMFPLNFNFTLNNKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
 DB 1 DVATWPHYLTADIRFCMFPLNFNFTLNNKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
 QY 61 IDPSIEIICRGYKGLLDGSDYPTAAYRKAAGVLRFGSGVIGNSRVKLTNVXDQVXR 120
 DB 61 IDPSIEIICRGYKGLLDGSDYPTAAYRKAAGVLRFGSGVIGNSRVKLTNVXDQVXR 120
 QY 121 GLVKEGEDQKVAADOLVNDGVDTLHTIGDDTNTAAALAPLANNNGLYTIGLPTV 180
 DB 121 GLVKEGEDQKVAADOLVNDGVDTLHTIGDDTNTAAALAPLANNNGLYTIGLPTV 180
 QY 181 DNDVFPKISLGAWTAAEGGARYFMNVVANNANPRLIVHEVMGRNGMLTAATQOER 240
 DB 181 DNDVFPKISLGAWTAAEGGARYFMNVVANNANPRLIVHEVMGRNGMLTAATQOER 240
 QY 241 KLLDRAEWLPEIGLTRESYEVHAFVPEMAIDLEAEAKRLREVMDKVCVNI FVSEGAV 300
 DB 241 KLLDRAEWLPEIGLTRESYEVHAFVPEMAIDLEAEAKRLREVMDKVCVNI FVSEGAV 300
 QY 301 EAIIVAEWQAKGEVPRDAGHITKLDVANGKWFGEPAOMIGAEKTLVQKSGYFAPASAS 360
 DB 301 EAIIVAEWQAKGEVPRDAGHITKLDVANGKWFGEPAOMIGAEKTLVQKSGYFAPASAS 360
 QY 361 NVDDMKRLISCADLAVECAFRRSGVIGHDEBNGVLRALIEPRIGKGFENIDTWMFNS 420
 DB 361 NVDDMKRLISCADLAVECAFRRSGVIGHDEBNGVLRALIEPRIGKGFENIDTWMFNS 420
 QY 421 MLSEIQPKGKVEVSH 437
 DB 421 MLSEIQPKGKVEVSH 437

RESULT 5
 US-08-416-870C-4
 Sequence 4, Application US/08416870C
 Patent No. 5824862

GENERAL INFORMATION:
 APPLICANT: HIYOSHI, TORU
 APPLICANT: MINE, TOSHIKI
 APPLICANT: KASAOKA, KEISUKE
 APPLICANT: TYSON, ROBERT HUM
 APPLICANT: PAGE, ANTHONY MILES JOHN
 TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 STREET: PO BOX 747
 CITY: FALL CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,870C
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MORPHY JR, GERALD M
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 760-195P (PCT)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 483 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-870C-4

Query Match 11.8%; Score 270; DB 2; Length 483;
 Best Local Similarity 27.6%; Pred. No. 16-17;
 Matches 113; Conservative 53; Mismatches 175; Indels 68; Gaps 15;

QY 36 ALTNAGGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSDYPTAAYRKAAGV 95
 DB 95 ALVTCGGLCPGLNTVIREIVCALYHMYGVTKVLGIDGGRGFSKVTITLTPKV--VND 151
 QY 96 LDRFGSVIGNSRVKLTNVKDQVXKGLVKEGEDQKVAADOLVNDGVDTLHTIGDDTNT 155
 DB 152 IHRGGTIIIGTSR-----GGHDKPKI-VDSIODRGINOVYIIIGSGTOK 194
 QY 156 AAADLAFLARNNGYGLTVGLPKTVNDVFPYKOSIGAWTAAEGGARYFMNVVANNANP 215
 DB 195 GAIVTYQEVRRRGKRAVAVGIRKTIIDNDIPVDKSGFPTVAEAGRAINAHVEESAE 254
 QY 216 RMLIVHEVMGRNGMLTAATQOERYKLDRAEWLPEIGLTRESYEVHAFVPEMAIDLEA 275
 DB 255 NGIGVVKLMGRYSGFIA-----MYATLASRDVDTCLIPESPFYLEG 295
 QY 276 EA-----KRLREVMDKVCVNI FVSEGAVEAIVAEQAGQEVPRDAPGHITKLDVANG 328
 DB 296 EGGLEBYEYERKLDQGHNV---IIVAEAGQGLLAELAEVK--TSTAQASNGKLLHDV- 348
 QY 329 PGKWFGEQ---FAOMIGAEKTLVQ-KSGYFAPASASNVDDMKRLISCADLAVECAFRR 383
 DB 349 -GLWISDKIKAFKIPKIPITLKYIDPTYIMRAVPSNASDN---VYCTLLAQSCYHGM 404
 QY 384 SGVIGHDE--NGNVLRAIEPRIGKGFENIDT--WNSMLSEIQP 428
 DB 405 AGVTGFTSLVNGR-QTYIPFRIRITEKQNNVVITDRMMAR-LLSSTNP 451

RESULT 6
 US-08-416-870C-6
 Sequence 6, Application US/08416870C
 Patent No. 5824862

GENERAL INFORMATION:
 APPLICANT: HIYOSHI, TORU
 APPLICANT: MINE, TOSHIKI
 APPLICANT: KASAOKA, KEISUKE
 APPLICANT: TYSON, ROBERT HUM
 APPLICANT: PAGE, ANTHONY MILES JOHN
 TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 STREET: PO BOX 747
 CITY: FALL CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22040-0747

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,870C
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 760-195P (PCT)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-870C-6

Query Match 11.6%; Score 264.5; DB 2; Length 469;
Best Local Similarity 26.8%; Pred. No. 3,4e-17;
Matches 109; Conservative 56; Mismatches 173; Indels 69; Gaps 16;

37 ILTAGLAAPCLNSAIGSLIERYTEIDPSIEICRGVYGLLLGDSYPTAEVRKAGVL 96
74 IYTCGGLCPGLNTVIREIVCGLNDYGVSRVIGCGVGFACWTIDSP---KSVNDN 130
97 QRFSGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQLVKDVDIHTIGDDPTNA 156
131 HRRGGTIVIGTSR-----GHDITMKI-VDSIODRGINQYVIGGGGTORG 173
157 AADLAFLARNNYGLTVIGLPTVNDVPIKOSIGAMTAAGARFENNVAAENNANR 216
174 AVITETIEIRRLKVAIVAGIPTIDNDIVIDRSFGFDVAEBQRAINAAHVEAGSAEN 233
217 MLIVHVGNGRCGMWLTAAATAOEYRKLDRBEMLPBGLTRESVEYHVAVFPEMAIDLEAE 276
234 GIGVVKLMNRHSGFLA---HYATLASR-----DVDCCIIPESPYLBEE 274
277 A-----KRLREVNDKYDCVNI FVSEGAIVEAIAEQAKGEVPRDAFGH- IKLDVAVN 328
275 GGLFYLEKRLKENGHMV---IIVAEGAGOKLINETKESMG---KDSAGNSILLDV-- 324
329 PKMFGGEORFAGMIGAEKTLVQ---KSGYFARASASNVDMRLIKSCADLAVECAFRRRS 384
325 -GLWISQKIKERPKKIKITINLKVIDPTVMIRAIIPNASDNYCTLLAHRVVHGMAAGYT 383
385 G-VIGHDEDNGVNLRAIEFPRI--KGGKPFNIDTDFNSMLSEIGQP 428
384 GTTVG--QVNGRHC-VIIPFRIITEKONKYSITDRMMAR-LISSTNOP 426

RESULT 7
US-08-416-870C-2
Sequence 2, Application US/08416870C
Patent No. 5824862
GENERAL INFORMATION:
APPLICANT: HIYOSHI, TORU
APPLICANT: MINE, TOSHIKI
APPLICANT: KASAKURA, KEISUKE
APPLICANT: TYSON, ROBERT HUM
APPLICANT: PAGE, ANTHONY WILES JOHN
TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP

STREET: PO BOX 747
CITY: FALL CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,870C
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 760-195P (PCT)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-870C-2

Query Match 11.5%; Score 262; DB 2; Length 485;
Best Local Similarity 25.7%; Pred. No. 6,2e-17;
Matches 113; Conservative 56; Mismatches 179; Indels 92; Gaps 17;

7 YHLTADIRFCWFLNFNFTLTKMKPKVAIITAGLAAPCLNSAIGSLIERYTEIDPSIE 66
86 YFSSDVAAC-----IVTCGGLCPGLNTVIREIVHSLDNYGVNKK 125
67 IICRGYKGLLLGDSYPTAEVRKAGVLDQRFSGSVIGNSRVKLTNVKDCVKGGLVEG 126
126 VFGIDGGYRGYFSKNIINLTP---KTVNDIHRGGTIISSR-----GG 166
127 EDPQKVAADQLVKDVDIHTIGDDPTNTAAADLAFLARNNYGLTVIGLPTVNDVFP 186
167 HDTTKI-VDSIODREINQYIITGGDTOKGAIVYIEIRRLKAKIYAGIPTIDNDIPV 225
187 IKOSIGAMTAAGARFENNVAAENNANPRMLIVHVGNGRCGMWLTAAATAOEYRKLDR 246
226 IDKSGFDTVAEBQRAINAAHVEASANGIGVVKLMGRYSGFIA----- 271
247 EMLPELIGTRESYEVHVAVFPEMAIDLEAE-----KRLREVNDKYDCVNI FVSEGA 299
272 ----MVAITLASRDVLDCLIPESPFYLEGDGLFEYIEKRLKENGHMV---IIVAEGAG 322
300 VEAI VAE--MOAKGEVPRDAFGH IKLDVAVNPGKMFGEQFAGMIGAEKTLVQ-----KSG 352
323 QELAEENAAHAKNEO---DASGNKLLQDV--GLWISQKIRHFP-ATKIKMPTTLKTIPT 376
353 YFARASASNVDMRLIKSCADLAVECAFRRRSVIGHDED--NGVNLRAIEFPRIKGGKP 410
377 YMIRAVPSNADSN---VYCTLLAQCCHVAGMAGYGFPSGLVNGR-QTYIPFNRIITEKON 432
411 FNIDTD--WFNSMLSEIGQP 428
433 MVYITDRMMAR-LISSTNOP 451

RESULT 8
US-09-902-540-16136
Sequence 16136, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 16136
LENGTH: 345
TYPE: PR
ORGANISM: Myxococcus xanthus
US-09-902-540-16136

Query Match 11.4%; Score 261; DB 4; Length 345;
Best Local Similarity 26.0%; Pred. No. 4.6e-17;
Matches 96; Conservative 66; Mismatches 141; Indels 66; Gaps 16;

QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLDGSDYPTAEVRKKA 93
DB 2 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLDGSDYPTAEVRKKA 93
QY 94 GYLORGGSGVIGNSRV---KLTIVKDCYKGLVKEGEDPOKAAADLVKDDVILHTTGG 150
DB 58 GLHR-GGTTILGTSRVNPFKEVGLERVKRA-----IERNGHVAVIAIG 101
QY 151 DDTNTAAADLAFLANNYGLTVIGLPTVNDVPPIKOSLGAWTAEGARF--MNV 208
DB 102 EGTLSAATM-----SQEGLRVGPVKTIIDNDINATDFPGDTVAIATEIDRLHST 155
QY 209 AENNANPMLIVHEVNGRCWMLTAATAOEYKRLDRAEWLPELGLTRESYEVHAFVE 268
DB 156 AESH---KXVIECVGHRVGMAT---YAGIAGAD-----VILVE 192
QY 269 MAIDLEAEAKRL--REVMKDVDCNIFVSEGAGVEAIVAEOMAKGQEVPRDAFGHKLDA 326
DB 193 IPADLAKVHEHIOHRAAGSTFESI-VVAEGTRIKLSADQOEQLVTSGALDEAGRRLG 251
QY 327 VNGKMGFQFQOMIGAETLVQSGYFAPASASNVDMRL-----IKSCADLAVBCAPR 381
DB 252 V--GTLIAEIERRTFE--TRVSVLGHIOHGAFTAHDRVLAIRYGVHAC--DMVANGERF 307
QY 382 RESGVIGHD 390
DB 308 KMAALRGND 316

RESULT 9
US-08-416-870C-10
Sequence 10, Application US/08416870C
Patent No. 5824862
GENERAL INFORMATION:
APPLICANT: HIYOSHI, TORU
APPLICANT: MINE, TOSHIKI
APPLICANT: KASAKURA, KEISUKE
APPLICANT: TYSON, ROBERT HUM
APPLICANT: PAGE, ANTHONY MILES JOHN
TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDE
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
STREET: PO BOX 747
CITY: PAUL CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,870C
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 760-195P (PCT)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-870C-10

Query Match 11.4%; Score 260; DB 2; Length 426;
Best Local Similarity 25.6%; Pred. No. 8e-17;
Matches 107; Conservative 65; Mismatches 174; Indels 72; Gaps 16;

QY 37 ILTAGGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLDGSDYPTAEVRKKA 96
DB 50 ITCGGLCPGLNTVIREIVGSLSYMGVKKIIGIEGYSGFYARNITDLDL--KTVNDI 106
QY 97 QRFSGSVINSKVKLTNVDCYKGLVKEGEDPOKAAADLVKDDVILHTTGGDDTNTA 156
DB 107 HRRGGTILGTSR-----GSHDTTKI-VDSIDRGINOVYIIGGDSOKG 149
QY 157 AADLAFLARNYGLTVIGLPTVNDVPPIKOSLGAWTAEGARF--MNV 216
DB 150 AAVIPEIRRRGKAVAGIPKTIIDNDIPITRSFGDTVAEAOAIAAHVATSFEN 209
QY 217 MEIVHEVNGRCWMLTAATAOEYKRLDRAEWLPELGLTRESYEVHAFVEPEMAIDLEAE 276
DB 210 GIGLVKLMGRVSGFLA-----MYATLASRDVDCCLIPSPFFLEGG 250
QY 277 -----AKRLAEVNDKVCNIFVSEGAGVEAIVAEOMAKGQEVPRDAFGHKLDA 329
DB 251 GGLFEFIRGLRKEIGHV---IYIAEGAG--ODLAESENGSTTL-KDASGNKLLQDV-- 302
QY 330 GKMFQFQFQOMIGAETLVQX---SGYFAPASASNVDMRLIKSCADLAVECAFRESGV 386
DB 303 GLMISORIKDHPAKKMTLTKYITDPTMIRAVPSNAD---NVCTTLAQSAYH---GV 355
QY 387 I-GHDEDNQNVLA---IEPRI--KGGKPNIDTDWFSNMLSEIGQKGGKVEVSH 437
DB 356 MAGYNGFTVGLVNGRHTYIPFYRIITEKQKVVITBMMAR-LISSTNGSPFMKHDH 412

RESULT 10
US-09-489-039A-12258
Sequence 12258, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12258
LENGTH: 338
TYPE: PR
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12258

CURRENT APPLICATION NUMBER: US/09/583,110
 CURRENT FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US/09/107,433
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US 60/085,131
 PRIOR FILING DATE: 1998-05-12
 PRIOR APPLICATION NUMBER: US 60/051,553
 PRIOR FILING DATE: 1997-07-02
 NUMBER OF SEQ ID NOS: 5322
 SEQ ID NO 5099
 LENGTH: 335
 TYPE: PR1
 ORGANISM: Streptococcus pneumoniae
 US-09-583-110-5099

Query Match 9.6%; Score 220; DB 4; Length 335;
 Best Local Similarity 25.9%; Pred. No. 4,7e-13;
 Matches 88; Conservative 55; Mismatches 115; Indels 82; Gaps 16;

QY 33 KVAAILTAGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSDYPTAEVRKK 92
 DB 2 KRIAVLTSGDAPGMAAIRAVRQ--AISEGMEVFGIYDYGAVAGSHPLDA---AS 56
 QY 93 AGVLPFGSGVIGNSR---VKLTNVKDCVKGGLVEGEDPOKVAADOLVKQGVDTLHTI 148
 DB 57 VGDIIIRGGTFPLHSARYPEFAQL-----EG---QLKGIEQLKKIGIEGVVVI 100
 QY 149 GDDDTTAADLAFLARNNYGLTVIGLPTVNDVFPKISIGATAEAGARFMNV 208
 DB 101 GGDGTHGAMRL-----TEHGPAIGLPTDNDVGTDFPTGPTAVTT--AMDALNKI 153
 QY 209 AENNAIPMLIVHEVGRNGC---WLTATAQEVKLLDRAEWLDELGLTRE----- 257
 DB 154 RRTSSHRTPFTEVGRNAGDIALWAGLATGD--EII-----IEAGFKMEDIVASTIK 206
 QY 258 -SYEV--HAVVPEKAIDLEAE-AKRLREVDKDVCAIIVSEGAIVEAIWQAKQ 312
 DB 207 AGVECCGKKNIIIVLAEGVMSAAEFQCKLKEAGDTS---LRVTE-----LGHIGRGS 256
 QY 313 EVPRDAFGHIKLDVAVNGKMFGEQFQMGIAEKTIVQKSG 352
 DB 257 PTARD-----RVLASRMGAHAIVKLLKEG 279

RESULT 14
 US-09-134-000C-5332
 Sequence 5332, Application US/09134000C
 Patent No. 6617156
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 032796-032
 CURRENT APPLICATION NUMBER: US/09/134,000C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/055,778
 PRIOR FILING DATE: 1997-08-15
 NUMBER OF SEQ ID NOS: 6812
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 5332
 LENGTH: 321
 TYPE: PR1
 ORGANISM: Enterococcus faecalis
 US-09-134-000C-5332

Query Match 9.6%; Score 218.5; DB 4; Length 321;
 Best Local Similarity 24.5%; Pred. No. 6.2e-13;
 Matches 92; Conservative 54; Mismatches 113; Indels 117; Gaps 15;

QY 33 KVAAILTAGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGSDYPTAEVRKK 92
 DB 3 KRIAVLTSGDAPGMAAIRAVRQ--AISEGMEVFGIYDYGAVAGSHPLDA---AS 56

QY 93 AGVLPFGSGVIGNSRVLKLTNVKDCVKG-----LVKEGEDPOKVAADOLVKD 140
 DB 52 -----RLVDVADVGDKIQRGTFLYSARYPEFATBEQ---QLKGIEQLKKF 93
 QY 141 GVDIHTTGDDTTRAAADLAFLARNNYGLTVIGLPTVNDVFPKISLGAMWTAEBG 200
 DB 94 GIEGLVIGGGDSYHGA-----MLTKRGPFAVGIPTIDNDIPCTDFTIGPDTA--- 143
 QY 201 ARYFMNVVAEN-----NANPMLIVHEVGRNGCMITAAVQERYKLLDRAEWLPEL 252
 DB 144 ---INTVLESIDRIIDRTATSHVTFVT--EVNGRNG-----DIALMSGVA 184
 QY 253 GLTRSEYEVHAFVPEMAIDLEAEAKRLREVD--KVDCAIIVSEGAIVEAIWQAK 310
 DB 185 GGADE-----IIEPHDFDMGNVAKRIEGRDRKCHCL--IILAEG----- 224
 QY 311 GGEVPRDAFGHIKLDVAVNGKMFGEQFQMGIAEKTIVQKSGIFARASASNDMLIKS 370
 DB 225 -----VWGNFADKLSE--YDFHTRVSIILGHVVRGSGPSARDVLAASK 267
 QY 371 CADLAVECAFRESGV 386
 DB 268 FGSYAVELLKEGKGL 283

RESULT 15
 US-09-107-532A-6470
 Sequence 6470, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSES: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 6470:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 323 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURES:
 NAME/KEY: m1bc_feature
 LOCATION: (B) LOCATION 1...323

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 14:22:36 ; Search time 133 Seconds
(without alignment)
1075.221 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTMPYHLTADIRCHWF.....FNSMLSEIGPKGKGVESH 437

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/PCTUS_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pcp.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 2284 | 100.0 | 437 | 9 | US-09-934-901-16 |
| 2 | 2284 | 100.0 | 437 | 9 | US-09-934-868-6 |
| 3 | 2284 | 100.0 | 437 | 10 | US-09-941-947A-2 |
| 4 | 2284 | 100.0 | 437 | 14 | US-10-320-924-16 |
| 5 | 2284 | 100.0 | 437 | 14 | US-10-320-874-16 |
| 6 | 2284 | 100.0 | 437 | 15 | US-10-363-567-2 |
| 7 | 2284 | 100.0 | 437 | 16 | US-10-321-210-16 |
| 8 | 659 | 28.9 | 184 | 15 | US-10-369-493-14348 |
| 9 | 306.5 | 13.4 | 350 | 14 | US-10-369-493-8108 |
| 10 | 294 | 12.9 | 342 | 14 | US-10-156-761-13610 |
| 11 | 293 | 12.8 | 414 | 15 | US-10-425-114-51700 |
| 12 | 293 | 12.8 | 525 | 15 | US-10-425-114-44512 |
| 13 | 293 | 12.8 | 527 | 15 | US-10-424-599-283989 |

| | | | | | | |
|----|-------|------|-----|----|----------------------|--------------------|
| 14 | 293 | 12.8 | 527 | 15 | US-10-425-114-49830 | Sequence 49830, A |
| 15 | 284.5 | 12.5 | 341 | 14 | US-10-156-761-14652 | Sequence 14652, A |
| 16 | 279.5 | 12.2 | 603 | 16 | US-10-437-963-204464 | Sequence 204464, A |
| 17 | 270 | 11.8 | 448 | 15 | US-10-282-122A-47251 | Sequence 47251, A |
| 18 | 269.5 | 11.8 | 341 | 14 | US-10-156-761-10359 | Sequence 10359, A |
| 19 | 269.5 | 11.8 | 530 | 16 | US-10-767-701-46418 | Sequence 46418, A |
| 20 | 263.5 | 11.5 | 769 | 16 | US-10-437-963-114757 | Sequence 114757, A |
| 21 | 261 | 11.4 | 345 | 15 | US-10-369-493-19261 | Sequence 19261, A |
| 22 | 259 | 11.3 | 541 | 16 | US-10-437-963-108835 | Sequence 108835, A |
| 23 | 255 | 11.2 | 434 | 15 | US-10-425-114-43593 | Sequence 43593, A |
| 24 | 254 | 11.1 | 507 | 15 | US-10-424-599-271108 | Sequence 271108, A |
| 25 | 253 | 11.1 | 319 | 15 | US-10-369-493-2905 | Sequence 2905, A |
| 26 | 250.5 | 11.0 | 357 | 15 | US-10-369-493-3669 | Sequence 3669, A |
| 27 | 248 | 10.9 | 447 | 15 | US-10-424-599-211067 | Sequence 211067, A |
| 28 | 246.5 | 10.8 | 987 | 15 | US-10-369-493-21998 | Sequence 21998, A |
| 29 | 246.5 | 10.8 | 987 | 16 | US-10-477-369-41 | Sequence 41, Appl |
| 30 | 241.5 | 10.6 | 428 | 15 | US-10-425-114-52834 | Sequence 52834, A |
| 31 | 241.5 | 10.6 | 481 | 15 | US-10-424-599-258809 | Sequence 258809, A |
| 32 | 240.5 | 10.5 | 320 | 9 | US-09-815-242-13785 | Sequence 13785, A |
| 33 | 240.5 | 10.5 | 320 | 15 | US-10-282-122A-72988 | Sequence 72988, A |
| 34 | 240.5 | 10.5 | 320 | 15 | US-10-282-122A-75033 | Sequence 75033, A |
| 35 | 240.5 | 10.5 | 320 | 15 | US-10-282-122A-76025 | Sequence 76025, A |
| 36 | 239.5 | 10.5 | 318 | 15 | US-10-369-493-9652 | Sequence 9652, A |
| 37 | 238.5 | 10.4 | 320 | 9 | US-09-815-242-10422 | Sequence 10422, A |
| 38 | 238.5 | 10.4 | 320 | 15 | US-10-369-493-23638 | Sequence 23638, A |
| 39 | 238.5 | 10.4 | 320 | 15 | US-10-282-122A-56771 | Sequence 56771, A |
| 40 | 235 | 10.3 | 352 | 15 | US-10-425-114-44609 | Sequence 44609, A |
| 41 | 234.5 | 10.3 | 327 | 15 | US-10-282-122A-78090 | Sequence 78090, A |
| 42 | 232 | 10.2 | 461 | 15 | US-10-282-122A-76398 | Sequence 76398, A |
| 43 | 229.5 | 10.0 | 295 | 15 | US-10-369-493-8855 | Sequence 8855, A |
| 44 | 229.5 | 10.0 | 320 | 15 | US-10-282-122A-56380 | Sequence 56380, A |
| 45 | 227.5 | 10.0 | 449 | 15 | US-10-369-493-8876 | Sequence 8876, A |

ALIGNMENTS

RESULT 1
US-09-934-901-16
Sequence 16, Application US/09934901
Patent No. US20020110885A1
GENERAL INFORMATION:
APPLICANT: Kofas, Matheos
APPLICANT: Odom, J. Martin
APPLICANT: No. US20020110885A1, Kelley C.
TITLE OF INVENTION: IDENTIFYING METHANOTROPHIC BACTERIAL STRAIN
FILE REFERENCE: CL619 US NA
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/229,906
PRIORITY FILING DATE: September 1, 2000
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Microsoft Office 97
SEQ ID NO 16
LENGTH: 437
TYPE: PRT
ORGANISM: METHYLIOMONAS SP.
US-09-934-901-16

| | | | | |
|-----------------------|--|--------------------|--------------|------------|
| Query Match | 100.0% | Score 2284 | DB 9 | Length 437 |
| Best Local Similarity | 100.0% | Pred. No. 4.8e-207 | | |
| Matches 437 | Conservative | 0 | Mismatches 0 | Indels 0 |
| Gap | | | | 0 |
| QY | 1 DVVTMPYHLTADIRCHWFPLNFYTLNKKKKAIAITLGGIAPCLNALSILIRYIE | 60 | | |
| DB | 1 DVVTMPYHLTADIRCHWFPLNFYTLNKKKKAIAITLGGIAPCLNALSILIRYIE | 60 | | |
| QY | 61 IDPSIEIICRGYKGLLGDSPYTAEVKKAAGVLOFGSGSVIGNSRKLTVNKDCVVR | 120 | | |
| DB | 61 IDPSIEIICRGYKGLLGDSPYTAEVKKAAGVLOFGSGSVIGNSRKLTVNKDCVVR | 120 | | |
| QY | 121 GLVKGEDPQKVAADQLVKDGVDIHTTIGGDDTNTAAADLAAFLANNYGLTVIGLPKTV | 180 | | |

DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPKQSLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAATAOEYR 240
DB 181 DNDVFPKQSLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRESYEYHAAVFPPEMAIDLEAARKLRVMDKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRESYEYHAAVFPPEMAIDLEAARKLRVMDKVCVNI FVSEGAGV 300
QY 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKMFGEQFAOMIGAETLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKMFGEQFAOMIGAETLVOKSGYFARASAS 360
QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVNLRAIEFPRIKGGKPFNIDTDFNS 420
DB 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVNLRAIEFPRIKGGKPFNIDTDFNS 420
QY 421 MLSEIGQPKGKVEYVSH 437
DB 421 MLSEIGQPKGKVEYVSH 437

RESULT 2

US-09-934-868-6
; Sequence 6, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 437
; TYPE: PRF
; ORGANISM: METHYLIOMONAS SP.
US-09-934-868-6

Query Match 100.0%; Score 2284; DB 9; Length 437;
Best Local Similarity 100.0%; Pred. No. 4,8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWMPYHLTADIRCHWFPLNFNFTYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWMPYHLTADIRCHWFPLNFNFTYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICYRGYKGLLDGSDYPVTAEVKKAGVLOFRGSGVIGNSRVKLTNVDCVVR 120
DB 61 IDPSIEIICYRGYKGLLDGSDYPVTAEVKKAGVLOFRGSGVIGNSRVKLTNVDCVVR 120
QY 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPKQSLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAATAOEYR 240
DB 181 DNDVFPKQSLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRESYEYHAAVFPPEMAIDLEAARKLRVMDKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRESYEYHAAVFPPEMAIDLEAARKLRVMDKVCVNI FVSEGAGV 300
QY 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKMFGEQFAOMIGAETLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKMFGEQFAOMIGAETLVOKSGYFARASAS 360

QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVNLRAIEFPRIKGGKPFNIDTDFNS 420
DB 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVNLRAIEFPRIKGGKPFNIDTDFNS 420
QY 421 MLSEIGQPKGKVEYVSH 437
DB 421 MLSEIGQPKGKVEYVSH 437

RESULT 3

US-09-941-947A-2
; Sequence 2, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzostowicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Koffas, Mattheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odum, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 US NA
; CURRENT APPLICATION NUMBER: US/09/941,947A
; PRIOR FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 437
; TYPE: PRF
; ORGANISM: Methyliomonas 16a
US-09-941-947A-2

Query Match 100.0%; Score 2284; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 4,8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVTWMPYHLTADIRCHWFPLNFNFTYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
DB 1 DVTWMPYHLTADIRCHWFPLNFNFTYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
QY 61 IDPSIEIICYRGYKGLLDGSDYPVTAEVKKAGVLOFRGSGVIGNSRVKLTNVDCVVR 120
DB 61 IDPSIEIICYRGYKGLLDGSDYPVTAEVKKAGVLOFRGSGVIGNSRVKLTNVDCVVR 120
QY 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
DB 121 GLVKGEDPQKVAADQVLDVGDVLIHTIGDDTNTAAADLAAFLARNNYGLTVIGLPTKV 180
QY 181 DNDVFPKQSLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAATAOEYR 240
DB 181 DNDVFPKQSLGAMTAAGAGARYFMNVVAENNANPRMLIVHEVMGRNCGMLTAATAOEYR 240
QY 241 KLLDRAEWLPELGLTRESYEYHAAVFPPEMAIDLEAARKLRVMDKVCVNI FVSEGAGV 300
DB 241 KLLDRAEWLPELGLTRESYEYHAAVFPPEMAIDLEAARKLRVMDKVCVNI FVSEGAGV 300
QY 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKMFGEQFAOMIGAETLVOKSGYFARASAS 360
DB 301 EAIIVAMQAKQGEVPRDAFGHIKLDVAVNPGKMFGEQFAOMIGAETLVOKSGYFARASAS 360
QY 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVNLRAIEFPRIKGGKPFNIDTDFNS 420
DB 361 NVDDMRILKSCADLAVECAFRRRESGVIGHDEBNGVNLRAIEFPRIKGGKPFNIDTDFNS 420
QY 421 MLSEIGQPKGKVEYVSH 437
DB 421 MLSEIGQPKGKVEYVSH 437

```
Db 421 MUSEIQPKGKVEVSH 437

RESULT 4
US-10-320-924-16
; Sequence 16, Application US/10320924
; Publication No. US20030129721A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Matheos
; APPLICANT: Odom, J. Martin
; APPLICANT: No. US20030129721A1ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/320,924
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-320-924-16

Query Match 100.0%; Score 2284; DB 14; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVTWPHYLTADIRFCHEFFLNFFYTLNKKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
Db 1 DVTWPHYLTADIRFCHEFFLNFFYTLNKKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
Qy 61 IDPSIEIICRYGGYKGLLDGSDYPTAEVRKKAQVLRFGSGVIGNSRYKLTVKDCVXR 120
Db 61 IDPSIEIICRYGGYKGLLDGSDYPTAEVRKKAQVLRFGSGVIGNSRYKLTVKDCVXR 120
Qy 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNYGLTVIGLPKTV 180
Db 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNYGLTVIGLPKTV 180
Qy 181 DNDVPEIKOSLGAMTAEQARFYEMNVVAENNANPMLLVHEVMGNCMLTAATQOER 240
Db 181 DNDVPEIKOSLGAMTAEQARFYEMNVVAENNANPMLLVHEVMGNCMLTAATQOER 240
Qy 241 KLLDRAEWLPELGLTRESYEVAHVPVEMAIIDLEAKRLREVMKVDVCNIFVSEGAGV 300
Db 241 KLLDRAEWLPELGLTRESYEVAHVPVEMAIIDLEAKRLREVMKVDVCNIFVSEGAGV 300
Qy 301 EAIIVAMQAKGOEVPDAFGHIKLDVNPCKMFGEOFAQMIQAEKTLVOKSGYFAPASAS 360
Db 301 EAIIVAMQAKGOEVPDAFGHIKLDVNPCKMFGEOFAQMIQAEKTLVOKSGYFAPASAS 360
Qy 361 NVDDMLIKSCADLAVECAFRESGYIGHEDNGNVLRAIEPPRIKGGKPFNIDTWFNS 420
Db 361 NVDDMLIKSCADLAVECAFRESGYIGHEDNGNVLRAIEPPRIKGGKPFNIDTWFNS 420
Qy 421 MUSEIQPKGKVEVSH 437
Db 421 MUSEIQPKGKVEVSH 437

RESULT 5
US-10-320-874-16
; Sequence 16, Application US/10320874
; Publication No. US20030138909A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Matheos
; APPLICANT: Odom, J. Martin
; APPLICANT: No. US20030138909A1ton, Kelley C.
```

```
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/320,874
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-320-874-16

Query Match 100.0%; Score 2284; DB 14; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVTWPHYLTADIRFCHEFFLNFFYTLNKKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
Db 1 DVTWPHYLTADIRFCHEFFLNFFYTLNKKPKVAIITAGGLAPCLNSAIGSLIRYTE 60
Qy 61 IDPSIEIICRYGGYKGLLDGSDYPTAEVRKKAQVLRFGSGVIGNSRYKLTVKDCVXR 120
Db 61 IDPSIEIICRYGGYKGLLDGSDYPTAEVRKKAQVLRFGSGVIGNSRYKLTVKDCVXR 120
Qy 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNYGLTVIGLPKTV 180
Db 121 GLVKEGEDPQKVAADLVKDGVDIILHTIGDDTNTAAADLAFLANNYGLTVIGLPKTV 180
Qy 181 DNDVPEIKOSLGAMTAEQARFYEMNVVAENNANPMLLVHEVMGNCMLTAATQOER 240
Db 181 DNDVPEIKOSLGAMTAEQARFYEMNVVAENNANPMLLVHEVMGNCMLTAATQOER 240
Qy 241 KLLDRAEWLPELGLTRESYEVAHVPVEMAIIDLEAKRLREVMKVDVCNIFVSEGAGV 300
Db 241 KLLDRAEWLPELGLTRESYEVAHVPVEMAIIDLEAKRLREVMKVDVCNIFVSEGAGV 300
Qy 301 EAIIVAMQAKGOEVPDAFGHIKLDVNPCKMFGEOFAQMIQAEKTLVOKSGYFAPASAS 360
Db 301 EAIIVAMQAKGOEVPDAFGHIKLDVNPCKMFGEOFAQMIQAEKTLVOKSGYFAPASAS 360
Qy 361 NVDDMLIKSCADLAVECAFRESGYIGHEDNGNVLRAIEPPRIKGGKPFNIDTWFNS 420
Db 361 NVDDMLIKSCADLAVECAFRESGYIGHEDNGNVLRAIEPPRIKGGKPFNIDTWFNS 420
Qy 421 MUSEIQPKGKVEVSH 437
Db 421 MUSEIQPKGKVEVSH 437

RESULT 6
US-10-363-567-2
; Sequence 2, Application US/10363567
; Publication No. US20040077068A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours & Company
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 PCT
; CURRENT APPLICATION NUMBER: US/10/363,567
; PRIOR FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: US 60/229858
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 60/229907
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 437
; TYPE: PRT
```

```

; ORGANISM: Methylobionas 16a
US-10-363-567-2

Query Match      100.0%; Score 2284; DB 15; Length 437;
Best Local Similarity 100.0%; Pred. No. 4,8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVTWTPYHLTADIRFCHEFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
Db 1 DVTWTPYHLTADIRFCHEFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
Qy 61 IDPSIEIICYRGYKGLLGSDSYPTAEVRKKAAGVLOFGSGSVIGNSRVKLTNNVDCVVR 120
Db 61 IDPSIEIICYRGYKGLLGSDSYPTAEVRKKAAGVLOFGSGSVIGNSRVKLTNNVDCVVR 120
Qy 121 GLVKEGEDPOKVAADOLVKDGVDLHTIGDDDTNTAAADLAFLARNNGYLVIGLPKTV 180
Db 121 GLVKEGEDPOKVAADOLVKDGVDLHTIGDDDTNTAAADLAFLARNNGYLVIGLPKTV 180
Qy 181 DNDVFPKISLGAMWTAEGARYPFNNVVAENNANPRMLIVHEVMGRNCGLTAAATQOYR 240
Db 181 DNDVFPKISLGAMWTAEGARYPFNNVVAENNANPRMLIVHEVMGRNCGLTAAATQOYR 240
Qy 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSBGAGV 300
Db 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSBGAGV 300
Qy 301 EAIIVAMQAKQOEVRDPAFGHI KLDVNPCKMFGSEQFQOMIGAETLVQKSGYFARASAS 360
Db 301 EAIIVAMQAKQOEVRDPAFGHI KLDVNPCKMFGSEQFQOMIGAETLVQKSGYFARASAS 360
Qy 361 NVDDMRLLIKSCADLAVECAFRRSGVIGHDEBNGVLAIEFPRIKGGKPFNIDTDWENS 420
Db 361 NVDDMRLLIKSCADLAVECAFRRSGVIGHDEBNGVLAIEFPRIKGGKPFNIDTDWENS 420
Qy 421 MLSEIGQPKGKVEVSH 437
Db 421 MLSEIGQPKGKVEVSH 437

RESULT 7
US-10-321-210-16
; Sequence 16, Application US/10321210
; Publication No. US20040115657A1
; GENERAL INFORMATION:
; APPLICANT: Kofeas, Matcheos
; APPLICANT: Odom, J. Martin
; APPLICANT: Norton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL619 US NA
; CURRENT APPLICATION NUMBER: US/10/321,210
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-321-210-16

Query Match      100.0%; Score 2284; DB 16; Length 437;
Best Local Similarity 100.0%; Pred. No. 4,8e-207;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVTWTPYHLTADIRFCHEFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
Db 1 DVTWTPYHLTADIRFCHEFFLNFPYTLNKKPKVAIITAGGLAPCLNSAIGSLIERYTE 60
```

```

Qy 61 IDPSIEIICYRGYKGLLGSDSYPTAEVRKKAAGVLOFGSGSVIGNSRVKLTNNVDCVVR 120
Db 61 IDPSIEIICYRGYKGLLGSDSYPTAEVRKKAAGVLOFGSGSVIGNSRVKLTNNVDCVVR 120
Qy 121 GLVKEGEDPOKVAADOLVKDGVDLHTIGDDDTNTAAADLAFLARNNGYLVIGLPKTV 180
Db 121 GLVKEGEDPOKVAADOLVKDGVDLHTIGDDDTNTAAADLAFLARNNGYLVIGLPKTV 180
Qy 181 DNDVFPKISLGAMWTAEGARYPFNNVVAENNANPRMLIVHEVMGRNCGLTAAATQOYR 240
Db 181 DNDVFPKISLGAMWTAEGARYPFNNVVAENNANPRMLIVHEVMGRNCGLTAAATQOYR 240
Qy 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSBGAGV 300
Db 241 KLLDRAEWLPELGLTRESYEVAHVPEMAIDLEAARLREVMKVCVNI FVSBGAGV 300
Qy 301 EAIIVAMQAKQOEVRDPAFGHI KLDVNPCKMFGSEQFQOMIGAETLVQKSGYFARASAS 360
Db 301 EAIIVAMQAKQOEVRDPAFGHI KLDVNPCKMFGSEQFQOMIGAETLVQKSGYFARASAS 360
Qy 361 NVDDMRLLIKSCADLAVECAFRRSGVIGHDEBNGVLAIEFPRIKGGKPFNIDTDWENS 420
Db 361 NVDDMRLLIKSCADLAVECAFRRSGVIGHDEBNGVLAIEFPRIKGGKPFNIDTDWENS 420
Qy 421 MLSEIGQPKGKVEVSH 437
Db 421 MLSEIGQPKGKVEVSH 437

RESULT 8
US-10-369-493-14348
; Sequence 14348, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14348
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14348

Query Match      28.9%; Score 659; DB 15; Length 184;
Best Local Similarity 67.9%; Pred. No. 6,9e-54;
Matches 125; Conservative 27; Mismatches 32; Indels 0; Gaps 0;

Qy 33 KKVAILTAGGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLGSDSYPTAEVRK 92
Db 1 QKVAILTAGGLAPCLNSAIGSLIERYSIDIAPEIDIAVRSYGVLGERIITIDMEX 60
Qy 93 AGVLOFGSGSVIGNSRVKLTNNVDCVVRKGLVKEGEDPOKVAADOLVKDGVDLHTIGDD 152
Db 61 AHLHRYGSGSPIGNSRVKLTNNVDCVVRKGLVKEGDNPRVAERLAADGIRITLHTIGDD 120
Qy 121 TTTAAADLAFLARNNGYLVIGLPKTVDNDVFPKISLGAMWTAEGARYPFNNVVAENN 212
Db 121 TTTAAADLAFLARNNGYLVIGLPKTVDNDVFPKISLGAMWTAEGARYPFNNVVAENN 212
Qy 213 ANPR 216
Db 181 AAPK 184
```

```

RESULT 9
US-10-369-493-8108
; Sequence 8108, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8108
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Thermobifida fusca
US-10-369-493-8108

Query Match      13.4%; Score 306.5; DB 15; Length 350;
Best Local Similarity 26.5%; Pred. No. 3.8e-20;
Matches 104; Conservative 63; Mismatches 145; Indels 81; Gaps 14;

QY 23 FNFYTLNKKRKAVALTAGGLAPCLNSAIGSLERYEIDPSIITICRGYKGLIGDS 82
DB 4 FKEVTOM---RVGLTGGDDCPGLNAVIRAVRKGIK-EYGEYFVGFGRDGMGPLEGDT 58
QY 83 YPTAERKKAQVLORFSGSVIGNSRVKLTNVKDCVRKGLVKEGEDPQKVAADQVLDVGV 142
DB 59 MPUDIEVLR--GILPR-GGITLSSRTNLMKIBGCVR-----VKDMMALGV 103
QY 143 DILHTTGGDDTNTAAADLAFLANNYGLTVGLPKTVNDVPEIKOSLGAMTAEGGAR 202
DB 104 DALVAIGEDDTLGVARQL-----HDGVMVVGVPKTIIDNLTADTVTFGDTAVNATE 157
QY 203 YFNNVVAENNAPRMILVHEVMGNCMLT-----AATQBYRKLDRAMELPELGITRE 257
DB 158 AIDRLHTTASSHRALTIV-EVMGRHAGMIALHAGMAGA----- 195
QY 258 SYEVHNAVFEPMADLLEAKRLREVMNDKVDVNIFFVSEGAGVEAIVAEWQAGQEVPRD 317
DB 196 ----NVLLIFRPPDIDVVAAYIESRKTNYAPITVVAEGHPRK---EGQLTLASGERD 247
QY 318 AFGHIKLDVAVPGKMFGEOPAQMIGAEKTLVQKS---GYPARASASVNDMLRIKSCADL 374
DB 248 SPGHVRLGGI-----GQRLAELEARTGKEARSVLVGHQRGSTPSAFPRVLTATRLGLH 301
QY 375 AVECAFRRESGVIGHDEDNQNV--LRAIEFPRI 405
DB 302 AITAV-----HDKDFGMVALRGTEIAYRV 325

RESULT 10
US-10-156-761-13610
; Sequence 13610, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761

```

```

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13610
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13610

Query Match      12.8%; Score 294; DB 14; Length 342;
Best Local Similarity 26.5%; Pred. No. 5.7e-19;
Matches 101; Conservative 66; Mismatches 140; Indels 74; Gaps 17;

QY 34 KVALITRAGGLAPCLNSAIGSLERYEIDPSIITICRGYKGLIGDSYPTAERKKA 93
DB 2 RVGLTGGDDCPGLNAVIRGVARKVQ-EYGYDFVGFGRDGMGPLEGDA--VRLDIPAVR 58
QY 94 GVLORFGGSVIGNSRVKLTNVKDCVRKGLVKEGEDPQKVAADQVLDVGLHTIGGDDT 153
DB 59 GILPR-GGITLSSRTNPLKLDGIR--IKE-----NLAKQEVDAIATIGEDT 105
QY 154 NTAADLAFLANNYGLTVGLPKTVNDVPEIKOSLGAMTAEGGARF--MNVVAEN 211
DB 106 LGVAARLT-----DEYGPVVGVPKTIIDNLTADTVTFGDTAVGATEAIDRLHTAES 160
QY 212 NANPRMLVHEVMGRCNGMTAATQBYRKLDRAMELPELGITRESYEVHNAVFEPMAL 271
DB 161 HMR--VLVEVGVGRHAGMTAHS-----GLAGGA--NVLLIFRPP 197
QY 272 DLEAKARLEREVMNDKVDVNIFFVSEGAGVEAIVAEWQAGQEVPR---DAFGHIKLDVAV 327
DB 198 DVUQVCAVYTSRKAASYAPITVVAEGA-----MPKQGMPLKDESLDSFGHRLSGV 249
QY 328 NPKWFGGEOPAQMIGAE-KTLVQKSGYFAPASASVNDMLRIKSCADLAVECAFRRESGV 386
DB 250 --GEMLAKIEKRTGKEARFTV--LGHVQRGSTPSAFDRVLTARFGIHAIEAV----- 298
QY 387 IGHDEDNQNV--LRAIEFPRI 405
DB 299 --RDGDFGMVALRGTDIAYRV 317

RESULT 11
US-10-425-114-51700
; Sequence 51700, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51700
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700837684_FLI.pep
US-10-425-114-51700

Query Match      12.8%; Score 293; DB 15; Length 414;
Best Local Similarity 26.8%; Pred. No. 9.3e-19;
Matches 113; Conservative 61; Mismatches 174; Indels 74; Gaps 15;

```


QY 327 VNPGRKFGQFAOMIGAETL-----VOKSGYFARASASNVDMRLIKSCADLAECAPR 381
 DB 377 V--GLWISQKIRSDHFAAQKTLPTLTKYIDPTMIRAIPEASNDVNYCTLLAQSAYHGAAM- 433
 QY 382 RESGVIGHDED--NGNVLAIEFPRIKGGKFPRIIDTD--WFSMSLSEIGOP-----KGK 432
 DB 434 --AGTYGYSGLVNGR-QTYIPFYRITERQNHVITDRMMAR-LTSTNQPSFLDAKGN 489
 QY 433 VE 434
 DB 490 EE 491

RESULT 14
 US-10-425-114-49830

/ Sequence 49830, Application US/10425114
 / Publication No. US20040034888A1
 / GENERAL INFORMATION:
 / APPLICANT: Liu, Jingdong
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kovalic, David K.
 / APPLICANT: Screen, Steven E.
 / APPLICANT: Tabaka, Jack E.
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 / FILE REFERENCE: 38-21(53313)B
 / CURRENT APPLICATION NUMBER: US/10/425,114
 / CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 73128
 / SEQ ID NO 49830
 / LENGTH: 527
 / TYPE: PRP
 / ORGANISM: Glycine max
 / FEATURE:
 / OTHER INFORMATION: Clone ID: 700559917_FLI.pep
 / US-10-425-114-49830

Query Match 12.8%; Score 293; DB 15; Length 527;
 Best Local Similarity 26.8%; Pred. No. 1.3e-18;

Matches 113; Conservative 61; Mismatches 174; Indels 74; Gaps 15;

QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGDSYPTAVERKKA 93
 DB 124 QAAITCGGLCPGLMIVIRVGLHMYGVKKVLGINSYRGFYARNITITLP---KGV 180
 QY 94 GVLQRFSGSVIGNSRVKLTNVKDCVKGGLVKGEDPOKVAADQLVKDGVDIHTTIGDDT 153
 DB 181 NDIHKRGCTVLTGSR-----GGHDTKKT-VDSIQDREGINQVYIIGDGT 223
 QY 154 NTPAADLAAFLARNNYGLTVIGLPTVDNDVPPIKQSLGAWTPAEGAGYFMNVVAENNA 213
 DB 224 QKGAISLIEFEVRKRGKAVSVGIPKTIIDVIDPIDSFGDTAVEEAQRIINAAYEAS 283
 QY 214 NPMRLIVHEVMGRNGCMULTAATAQERYKLLDRAEMLPGLTRESYEVAAVPEMAIDL 273
 DB 284 VENGIGVAVLMGRNSGFI-----WYATLASDNDVCCCLIPESPFL 324
 QY 274 EAEA-----KRLREVMKQVDCVNI-FVSEGAVEAIVAEQAKGQEVPRDAFGHIKLLDA 326
 DB 325 EGFGLYEYIERKRLKENGHNV---IIVAEGAOELVSESVOGMSKQ---DASGNLFPD 377
 QY 327 VNPGRKFGQFAOMIGAETL-----VOKSGYFARASASNVDMRLIKSCADLAECAPR 381
 DB 378 V--GLWISQKIRSDHFAAQKTLPTLTKYIDPTMIRAIPEASNDVNYCTLLAQSAYHGAAM- 434
 QY 382 RESGVIGHDED--NGNVLAIEFPRIKGGKFPRIIDTD--WFSMSLSEIGOP-----KGK 432
 DB 435 --AGTYGYSGLVNGR-QTYIPFYRITERQNHVITDRMMAR-LTSTNQPSFLDAKGN 490
 QY 433 VE 434

DB 491 EE 492

RESULT 15

US-10-156-761-14652
 / Sequence 14652, Application US/10156761
 / Publication No. US20030119018A1
 / GENERAL INFORMATION:
 / APPLICANT: OMURA, SATOSHI
 / APPLICANT: IKEDA, HARUO
 / APPLICANT: ISHIKAWA, JUN
 / APPLICANT: HORIKAWA, HIROSHI
 / APPLICANT: SHIBA, TADAYOSHI
 / APPLICANT: SAKAKI, YOSHIYUKI
 / APPLICANT: HATTORI, MASAHIRA
 / TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 / FILE REFERENCE: 249-262
 / CURRENT APPLICATION NUMBER: US/10/156,761
 / CURRENT FILING DATE: 2002-05-29
 / PRIOR APPLICATION NUMBER: JP 2001-204089
 / PRIOR FILING DATE: 2001-05-30
 / PRIOR APPLICATION NUMBER: JP 2001-272697
 / PRIOR FILING DATE: 2001-08-02
 / NUMBER OF SEQ ID NOS: 15109
 / SEQ ID NO 14652
 / LENGTH: 341
 / TYPE: PRP
 / ORGANISM: Streptomyces avermectilis
 / US-10-156-761-14652

Query Match 12.5%; Score 284.5; DB 14; Length 341;
 Best Local Similarity 27.5%; Pred. No. 4.5e-18;

Matches 97; Conservative 62; Mismatches 143; Indels 51; Gaps 11;

QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICRGYKGLLDGDSYPTAVERKKA 93
 DB 2 RIGVLTSGDCGCLNAVIRSVHR-AVVDHGDVIGFRDGMKGLCECDY--LKLDDAVS 58
 QY 94 GVLQRFSGSVIGNSRVKLTNVKDCVKGGLVKGEDPOKVAADQLVKDGVDIHTTIGDDT 153
 DB 59 GILAR-GGTTIASSSRVQPAHLDDVER-----ARGHVAELGIDAIIPIGGBGT 105
 QY 154 NTPAADLAAFLARNNYGLTVIGLPTVDNDVPPIKQSLGAWTPAEGAGYFMNVVAENNA 213
 DB 106 LKAARLL-----SDAGLPVIGVPTKTIIDVIDVTFGFTAVGAVATEALDLKTTAAS 159
 QY 214 NPMRLIVHEVMGRNGCMULTAATAQERYKLLDRAEMLPGLTRESYEVAAVPEMAIDL 273
 DB 160 HORVLTIV-EVMGRHGTWALHSGM-----AAGAAIAIVPERPFDI 198
 QY 274 EAEKRLREVMKQVDCVNI-FV-SEGAVEAIVAEQAKGQEVPRDAFGHIKLLDAVNPGRK 332
 DB 199 EELAAKVGRRFEAGKRAIIVAAEGAKPPAGSMDF-----DEGKDYVGHERRPAGT--AAQ 252
 QY 333 FGEQFAOMIGAETLVOKSGYFARASASNVDMRLIKSCADLAECAPRREG 385
 DB 253 LSLLEBERLKGKARPV-IIGHVQGGTPTAYDRAVLATRFGMHVAEVAHVHGEFG 304

Search completed: February 24, 2005, 14:34:59
 Job time : 135 secs

THIS PAGE BLANK (USP)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 14:01:14 ; Search time 182 Seconds

(without alignment)
1229.553 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTMPYHLTADIRFCWHF.....FNSMLSEIGQPKGKGVESH 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|--------------|------------------------------|
| 1 | 1531.5 | 67.1 | 453 | Q7UES2 | Q7UES2 rhodospirillum rubrum |
| 2 | 1453 | 63.6 | 410 | Q9NGP6 | Q9NGP6 mastigamoeba |
| 3 | 1432 | 62.7 | 399 | Q6A8S9 | Q6A8S9 propionibacterium |
| 4 | 1314 | 57.5 | 403 | PPF PROFR | PPF PROFR |
| 5 | 1302 | 57.0 | 404 | Q92AK9 | Q92AK9 rhizobium |
| 6 | 1292.5 | 56.6 | 479 | Q8UDL0 | Q8UDL0 agrobacterium |
| 7 | 1292.5 | 56.6 | 514 | Q7CKV0 | Q7CKV0 agrobacterium |
| 8 | 311 | 13.6 | 342 | 1 K6P1 STRCO | 1 K6P1 STRCO |
| 9 | 305 | 13.4 | 346 | 1 PFP DICTH | 1 PFP DICTH |
| 10 | 298.5 | 13.1 | 341 | 1 PFP AMYME | 1 PFP AMYME |
| 11 | 294 | 12.9 | 342 | 2 Q82AH4 | 2 Q82AH4 streptomycete |
| 12 | 284.5 | 12.5 | 341 | 2 Q826Q3 | 2 Q826Q3 streptomycete |
| 13 | 282 | 12.3 | 488 | 2 Q9FVU5 | 2 Q9FVU5 streptomycete |
| 14 | 281.5 | 12.3 | 341 | 1 K6P3 STRCO | 1 K6P3 STRCO |
| 15 | 281.5 | 12.3 | 485 | 2 Q9CSJ7 | 2 Q9CSJ7 streptomycete |
| 16 | 279.5 | 12.2 | 358 | 2 Q6AGJ2 | 2 Q6AGJ2 leifsonia |
| 17 | 278.5 | 12.2 | 542 | 2 Q6LAN6 | 2 Q6LAN6 oryza sativa |
| 18 | 278 | 12.2 | 366 | 1 K6P2 CHOPF | 1 K6P2 CHOPF |
| 19 | 275 | 12.0 | 447 | 1 Q660E6 | 1 Q660E6 borrelia |
| 20 | 274.5 | 12.0 | 341 | 1 PFP AMYMD | 1 PFP AMYMD |
| 21 | 271.5 | 11.9 | 341 | 1 K6P2 STRCO | 1 K6P2 STRCO |
| 22 | 270 | 11.8 | 448 | 2 Q516F9 | 2 Q516F9 borrelia |
| 23 | 269.5 | 11.8 | 341 | 2 Q82AD3 | 2 Q82AD3 streptomycete |
| 24 | 266.5 | 11.7 | 319 | 2 Q74BH3 | 2 Q74BH3 geobacter |
| 25 | 266 | 11.6 | 489 | 2 Q94AA4 | 2 Q94AA4 arabidopsis |
| 26 | 263.5 | 11.5 | 436 | 2 Q276S1 | 2 Q276S1 entamoeba |
| 27 | 263.5 | 11.5 | 586 | 2 Q94U12 | 2 Q94U12 oryza sativa |
| 28 | 260.5 | 11.4 | 775 | 2 Q91G72 | 2 Q91G72 oryza sativa |
| 29 | 260 | 11.4 | 382 | 2 Q24812 | 2 Q24812 entamoeba |
| 30 | 259.5 | 11.4 | 444 | 2 Q72AD6 | 2 Q72AD6 desulfovibrio |
| 31 | 257.5 | 11.3 | 500 | 2 Q95T07 | 2 Q95T07 arabidopsis |

| | | | | | |
|----|-------|------|-----|--------------|---------------------|
| 32 | 256.5 | 11.2 | 567 | 2 Q65X97 | Q65X97 oryza sativa |
| 33 | 256.5 | 11.2 | 991 | 2 Q6FTX6 | Q6FTX6 candida |
| 34 | 255.5 | 11.2 | 319 | 1 K6P2 CLOTE | Q89022 clostridium |
| 35 | 253.5 | 11.1 | 987 | 2 Q758H0 | Q758H0 ashyia |
| 36 | 253 | 11.1 | 319 | 1 K6P2 THEM4 | Q9W52 thermococcus |
| 37 | 253 | 11.1 | 473 | 2 Q817L4 | Q817L4 arabidopsis |
| 38 | 252 | 11.0 | 459 | 2 Q8VU09 | Q8VU09 amycolatops |
| 39 | 252 | 11.0 | 487 | 2 Q15648 | Q15648 trypanosoma |
| 40 | 251 | 11.0 | 473 | 2 Q9M0F9 | Q9M0F9 arabidopsis |
| 41 | 250 | 10.9 | 426 | 2 Q61068 | Q61068 trichomonas |
| 42 | 248.5 | 10.9 | 437 | 2 Q27705 | Q27705 naegleria |
| 43 | 246.5 | 10.8 | 987 | 1 K6P1 YEAST | P16861 saccharomyc |
| 44 | 245.5 | 10.7 | 992 | 1 K6P1 KUTLA | Q03215 kluyveromyc |
| 45 | 245 | 10.7 | 324 | 2 Q8GNC1 | Q8GNC1 haemophilus |

ALIGNMENTS

| RESULT 1 | ID | Q7UES2 | PRELIMINARY; | PRT; | 453 AA. |
|-----------------------|--|-------------------------------------|----------------------|-----------------------------|------------------|
| AC | Q7UES2 | 01-OCT-2003 (TRMBLrel. 25, Created) | | | |
| DT | 01-OCT-2003 (TRMBLrel. 25, Last sequence update) | | | | |
| DT | 01-MAR-2004 (TRMBLrel. 26, Last annotation update) | | | | |
| DE | PPI-phosphofructokinase (EC 2.7.1.90) | | | | |
| GN | Name=Pfk; OrderedLocustNames=RB10591; | | | | |
| OS | Rhodospirillum rubrum | | | | |
| OC | Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales; | | | | |
| OC | Planctomycetaceae; Pirellula. | | | | |
| OX | NCBI_TaxID=117; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=1; | | | | |
| RX | MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100; | | | | |
| RA | Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T., | | | | |
| RA | Ludwig W., Gade A., Borzym K., Heilmann K., Rabus R., | | | | |
| RA | Schlesner H., Amann R., Reinhardt R.; | | | | |
| RT | "Complete genome sequence of the marine planctomycete Pirellula sp. | | | | |
| RT | strain 1." | | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003). | | | | |
| DR | EMBL; BX294151; CAD78962.1; - | | | | |
| DR | GO; GO:0005945; C:6-phosphofructokinase complex; IEA. | | | | |
| DR | GO; GO:0008772; F:6-phosphofructokinase activity; IEA. | | | | |
| DR | GO; GO:0047334; F:diphosphate-fructose-6-phosphate 1-phosphot. .; IEA. | | | | |
| DR | GO; GO:0016301; F:kinase activity; IEA. | | | | |
| DR | GO; GO:0016740; F:transferase activity; IEA. | | | | |
| DR | GO; GO:006096; P:glycolysis; IEA. | | | | |
| DR | InterPro; IPR000023; Pfkfructokinase. | | | | |
| DR | Pfam; PF00365; Pfk; 1. | | | | |
| DR | Prodom; PD000707; PpfFructokinase; 1. | | | | |
| KW | Complete proteome; Kinase; Transferase. | | | | |
| SEQ | SEQUENCE 453 AA; 49268 MW; 9D7D76C8C909DAB CRC64; | | | | |
| Query Match | 67.1%; Score 1531.5; DB 2; Length 453; | | | | |
| Best local similarity | 65.9%; Pred. No. 3.5e-88; | | | | |
| Matches | 286; Conservative 61; Mismatches 78; Indels 9; Gaps 1; | | | | |
| QY | 4 | TVPYHLTADIRFCWHF | FLNENFYTLNKKPKVA | ILTAGGLAPCLNSAIGSLIERYTEIDP | 63 |
| DB | 29 | SWKIHAFASLESEPLMSI | -----KRVGILTLAGLAPCL | LSAIGALIEAYEQAP | 79 |
| QY | 64 | STFETTCRGYKGLLDGSD | PYTAAYRKAAGLQRF | SGSVINSRVKTNNVDCVRRGIV | 123 |
| DB | 80 | ETETTCRSGYKGLLDGSD | FFVDAHSRKNAAILH | GGSPINSRVKLTNVADCVRRGIV | 139 |
| QY | 124 | KGSDPQKVAADQVADLV | GDVLTATIGDQDNTA | AAADLAFLARNNYGLTVIGL | PKTVND 183 |
| DB | 140 | SSGQDPLQVAARLQSD | EDVDVLTATIGDQDNT | TAADLAFLARNNYGLTVIGL | PKTVND 199 |
| QY | 184 | VPTIKOSLGAMTAAAG | QAGARYFMNVVVAEN | NANFRMLIVHEVMGRNCGL | TAAATQAEYRKL 243 |

DB 200 VIPITQSGANTAAEGAKFFENVYGERNNAPRMILIEHWGRNCGMLTAATAKTRRL 259
QY 244 DRAEWLPELGLTRRESYEVHVAVPEPMALDELEAKRLBEMVDKVCNIPVSEGAQVAI 303
DB 260 OKANFLPEMGJLSQBRDVGAVFVPEMSFPLEEAKRLRAIMEIDCVNIFISEGAGVDTI 319
QY 304 VAEMQAKQOEVRDPAFGHIKLDVNPGRKFGQPAQMTGAETLVQKSGYFAPARASVND 363
DB 320 VKEMSRGESVPEKDFGHYKLDVNPGRKFGQPAQMTGAETLVQKSGYFAPARASVND 379
QY 364 DMRLIKSCADLAVECAPRESGVIQHDNDGNVLAIEFPRIKGGKFPNIDTFNSMLS 423
DB 380 DIALIGRCAKKAQVEACMGQIGSVGEDDQNNELRAIEFERLAGSKFPDINWTFGLDLS 439
QY 424 EIGQPKGKXVEVSH 437
DB 440 QMGPKGKXVEVSH 453

RESULT 2

QNGP6 PRELIMINARY; PRT; 410 AA.
AC QNGP6; 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE PPI-phosphofructokinase (EC 2.7.1.11).
GN Name=Pp1pfk;
OS Maestigamoeba balamuchi (Phreatamoeba balamuchi).
OC Eukaryota; Pelobiontida; Maestigameobidae; Maestigameoba.
NCBI_Taxid=108607;
RX MEDLINE=21528850; PubMed=11673446;
RX DOI=10.1128/JB.183.22.6714-6716.2001;
RA Muller M., Lee J.A., Gordon P., Gaasterland T., Sensen C.W.:
RT "Presence of prokaryotic and eukaryotic species in all subgroups of
the Pp(1)-dependent group II phosphofructokinase protein family.";
RL J. Bacteriol. 183:6714-6716(2001).
DR EMBL; AF246209; AAF70463.1; -
DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
DR GO; GO:0016301; E:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR00023; Pfruckinase.
DR InterPro; IPR011405; PPI-PFK_SMC01852.
DR Pfam; PF00365; PFK, 1.
DR PIRSF; PIRSF016484; PPI-PFK_SMC01852; 1.
DR PRINTS; PR00476; PHERCTKINASE.
DR KINASE; Transferase.
SQ SEQUENCE 410 AA; 44200 MW; CD362D1D2D41A0D0 CRC64;

Query Match 63.6%; Score 1453; DB 2; Length 410;
Best Local Similarity 67.2%; Pred. No. 2, 7e-83;
Matches 272; Conservative 55; Mismatches 78; Indels 0; Gaps 0;

QY 33 KKVAILTGAAGLAPCLNSAIGSLIERYTEIDPSIEIICRYGGYKGLLDGSPVTAVERKK 92
DB 3 KTVALLTGGALPCLSSAVGLIERYTEISPEFTNIILVINGKGLLGEKVLVTAMMLQ 62
QY 93 AGVLQRFSGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQIVKQGVILHTIGDD 152
DB 63 AAVLHTVGGSCIGNSRVKMANVADCVKGLVEGQDPQVADQILIKQGVILHTIGDD 122
QY 153 TMTAAADLAALFALANNYGLTVIGPKTYNDVFPKIQSGAMTAEGQARFPMNVAAEN 212
DB 123 TMTAAADLAALFALANNYGLTVIGPKTYNDVFPKIQSGAMTAEGQARFPMNVAAEN 182
QY 213 ANPRLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLTRSEYEVHVAVPEMAID 272
DB 243 ANPRLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLTRSEYEVHVAVPEMAID 242

QY 273 LEAEAKRLREVMVDKVCNIPVSEGAQVVAEVAEMQAKQOEVRDPAFGHIKLDVNPGRK 332
DB 243 LKAAARLRVAMREVCNIPVSEGAQVVAEVAEMQAKQOEVRDPAFGHIKLDVNPGRK 302
QY 333 FGEQPAQMTGAETLVQKSGYFAPARASVNDMLIKSCADLAVECAPRESGVIQHDND 392
DB 303 FGEQPAQMTGAETLVQKSGYFAPARASVNDMLIKSCADLAVECAPRESGVIQHDND 362
QY 393 NGVLAIEFPRIKGGKFPNIDTFNSMLS EIGQPKGKXVEVSH 437
DB 363 NGVLAIEFPRIKGGKFPNIDTFNSMLS EIGQPKGKXVEVSH 407
RESULT 3
Q6A8S9 PRELIMINARY; PRT; 399 AA.
ID Q6A8S9
AC Q6A8S9; 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Pyrophosphate-fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90).
GN OrderedLocNames=PPA1090;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
NCBI_Taxid=1747;
RX SEQUENCE FROM N.A.
RX STRAIN=KPA173202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hostler F., Liesegang H., Wielez A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.:
RT "The complete genome sequence of Propionibacterium acnes, a commensal
of human skin";
RL Science 305:671-673(2004).
DR EMBL; AE017283; AAT82837.1; -
DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR00023; Pfruckinase.
DR Pfam; PF00365; PFK, 1.
DR PRINTS; PR00476; PHERCTKINASE.
DR Prodom; PD000707; Pfruckinase; 1.
DR KINASE; Transferase.
SQ SEQUENCE 399 AA; 42944 MW; A0A845D7FAB0C1CD CRC64;

Query Match 62.7%; Score 1432; DB 2; Length 399;
Best Local Similarity 68.4%; Pred. No. 5, 4e-82;
Matches 270; Conservative 54; Mismatches 71; Indels 0; Gaps 0;

QY 33 KKVAILTGAAGLAPCLNSAIGSLIERYTEIDPSIEIICRYGGYKGLLDGSPVTAVERKK 92
DB 4 KKVALLTGGAPCLSTHISGLIQRYTEVAPBEVILAKHGEGLKDDFLEVDTYVKN 63
QY 93 AGVLQRFSGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQIVKQGVILHTIGDD 152
DB 64 AEILKRFSGSVIGNSRVKLTNVKDCVKGGLVEGEDPQKVAADQIVKQGVILHTIGDD 123
QY 153 TMTAAADLAALFALANNYGLTVIGPKTYNDVFPKIQSGAMTAEGQARFPMNVAAEN 212
DB 124 TMTAAADLAALFALANNYGLTVIGPKTYNDVFPKIQSGAMTAEGQARFPMNVAAEN 183
QY 213 ANPRLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLTRSEYEVHVAVPEMAID 272
DB 184 SGRMLIYHEVWGRNCGMLTAATAOEYRKLDRAEWLPGLTRSEYEVHVAVPEMAID 243
QY 273 LEAEAKRLREVMVDKVCNIPVSEGAQVVAEVAEMQAKQOEVRDPAFGHIKLDVNPGRK 332
DB 244 LEAEAKRLREVMVDKVCNIPVSEGAQVVAEVAEMQAKQOEVRDPAFGHIKLDVNPGRK 303
QY 333 FGEQPAQMTGAETLVQKSGYFAPARASVNDMLIKSCADLAVECAPRESGVIQHDND 392

DB 304 FGEQFADKLGAEVWQKSGYFSRSASNSADLELLGRTCDLAVDCAAGKGVIGQDEE 363
 QY 393 NGNVLAIRAEFPRIKGGKPFNIDTWFNSMLSEIGQ 427
 DB 364 NGDITLNIAPDRIKGGKPFDTTQPMFTALMSEIGQ 398

RESULT 4
 PEP_PROFR
 ID_PEP_PROFR STANDARD; PRT; 403 AA.
 P29595;
 DT 01-APR-1993 (rel. 25, Created)
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 25-OCT-2004 (rel. 45, Last annotation update)
 DE Pyrophosphatase--fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90)
 DE (6-phosphofructokinase, pyrophosphate dependent) (Pyrophosphatase-
 DE dependent 6-phosphofructose-1-kinase) (PPI-PFK).
 DE Phosphofructokinase) (PPI-PFK).
 GN Name=ppf; Synonym=ppik;
 OS Propionibacterium freudenreichii shermanii.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
 OC NCBI_Taxid=1752;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91358443; PubMed=1653240;
 RA Lador U.S., Gollupudl L., Tripathi R.L., Latschew S.P., Kemp R.G.,
 RT "Cloning, sequencing, and expression of pyrophosphate-dependent
 RT phosphofructokinase from Propionibacterium freudenreichii.";
 RN J. Biol. Chem. 266:16550-16555(1991).
 RN [2]
 RP IDENTIFICATION OF CRITICAL LYSYL RESIDUES.
 RX MEDLINE=92273593; PubMed=1317210;
 RA Green P.C., Latschew S.P., Lador U.S., Kemp R.G.,
 RT "Identification of critical lysyl residues in the pyrophosphate-
 RT dependent phosphofructo-1-kinase of Propionibacterium
 RT freudenreichii.";
 RN Biochemistry 31:4815-4821(1992).
 CC -1- CATALYTIC ACTIVITY: Diphosphate + D-fructose 6-phosphate =
 CC phosphate + D-fructose 1,6-bisphosphate.
 CC -1- ENZYME REGULATION: Non-allosteric.
 CC -1- PATHWAY: Key control step of glycolysis.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: Belongs to the phosphofructokinase family. PPI
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M67447; AAA25675.1; -
 DR PIR; A4169; A4169.
 DR InterPro; IPR000023; PpfFructKinase.
 DR Pfam; PF00365; PFK; 1. PFK_SMC01852; 1.
 DR PIRSF; PIRSF036484; PPI-PFK_SMC01852; 1.
 DR PRINTS; PR00476; PPFCKTKINASE.
 DR ProDom; PD000707; PpfFructKinase; 1.
 KW ATP-binding; Direct protein sequencing; Kinase; Magnesium;
 KW Transference.
 FT INIT MET 0
 FT NP_BIND 21 25 ATP (By similarity).
 FT NP_BIND 120 125 Pyrophosphate (potential).
 FT NP_BIND 178 182 ATP (By similarity).
 FT NP_BIND 195 231 ATP (By similarity).
 FT ACT_SITE 150 150 Proton acceptor (By similarity).
 FT BINDING 325 325 Substrate (By similarity).
 SQ SEQUENCE 403 AA; 43114 MM; 1DEB3272B4A2B6E6 CRC64;

Query Match 57.5%; Score 1314; DB 1; Length 403;
 Best Local Similarity 62.6%; Pred. No. 1,4e-74;
 Matches 248; Conservative 60; Mismatches 88; Indels 0; Gaps 0;

QY 33 KVVAILTAGGAPCPUNSAIGSLIERYTEIDPSEIETCYGGYKGLLGDSYPTAEVRKK 92
 DB 2 KVVAILTAGGAPCPUNSAIGSLIERYTEIDPSEIETCYGGYKGLLGDSYPTAEVRKK 61
 QY 93 AGVLAGFGSGVIGNSRVKLTNYVDCVKGGLYKEGEDPQKVAADQVLKDGVDLIHTIGDD 152
 DB 62 YRLRFSFGSGPIGNSRVKLTNYVDCVKGGLYKEGEDPQKVAADQVLKDGVDLIHTIGDD 121
 QY 153 TWTAADLAFLARNNGYLTGTPKTVNDVFPFKQSIGAVTAEGARYFNVAENN 212
 DB 122 TWTAADLAFLARNNGYLTGTPKTVNDVFPFKQSIGAVTAEGARYFNVAENN 181
 QY 213 ANPRMLIVHNGRNCGWLTAAQCYRKLDBAEKLPBLGLTRESYEVHNVPEMAID 272
 DB 182 AAPRELIIHEINGRCNGYLAETSRRYVAMLAQQLPEAGLDRGMDIHALVPEATID 241
 QY 273 LBAEAKRLBEMDKYDCVNI FVSEGAGVEAIVEMQAKGQVEPRDAFGHIKLDVANPGKY 332
 DB 242 LBAEAKRLBEMDKYDCVNI FVSEGAGVEAIVEMQAKGQVEPRDAFGHIKLDVANPGKY 301
 QY 333 FGEQFADKLGAEVWQKSGYFSRSASNSADLELLGRTCDLAVDCAAGKGVIGQDEE 392
 DB 302 FGEQFADKLGAEVWQKSGYFSRSASNSADLELLGRTCDLAVDCAAGKGVIGQDEE 361
 QY 393 NGNVLAIRAEFPRIKGGKPFNIDTWFNSMLSEIGQ 428
 DB 362 NGDITLNIAPDRIKGGKPFDTTQPMFTALMSEIGQ 397

RESULT 5
 ID Q92NK9 PRELIMINARY; PRT; 404 AA.
 Q92NK9;
 AC Q92NK9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE PROBABLE PYROPHOSPHATE--FRUCTOSE 6-PHOSPHATE 1-PHOSPHOTRANSFERASE
 DE PROTEIN (EC 2.7.1.90).
 DE ORFNames=SMC01852;
 GN Rhizobium meliloti (Sinorhizobium meliloti).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OC NCBI_Taxid=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
 RA Capela D., Batory-Hubler F., Guzy J., Bothe G., Anpe F., Batut J.,
 RA Boisard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godt T., Goffeau A., Kahn D., Kiss E., Lelaire V., Masny D.,
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL; AL591789; CAC46769.1; -
 DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
 DR GO; GO:0047334; F:diphosphate-fructose-6-phosphate 1-phosphot. . .; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR CO; GO:006096; P:glycolysis; IEA.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000023; PpfFructKinase.
 DR InterPro; IPR011405; PPI-PFK_SMC01852.
 DR Pfam; PF00365; PFK; 1.
 DR PIRSF; PIRSF036484; PPI-PFK_SMC01852; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR PRINTS; PR00476; PPFCKTKINASE.
 DR ProDom; PD000707; PpfFructKinase; 1.

KW Complete proteome; Transferase.
 SQ SEQUENCE 404 AA; 43712 MW; 1E6E5CD966A13D5 CRC64;
 Query Match 57.0%; Score 1302; DB 2; Length 404;
 Best Local Similarity 62.3%; Pred. No. 8.1e-74;
 Matches 246; Conservative 59; Mismatches 90; Indels 0; Gaps 0;

QY 33 KKVAILTGLGLAPCLNSAIGSLIERYTEIDPSIEITICRGYKGLLGDSPVTAEVK 92
 DB 4 KKVAILTGLGLAPCLNSAIGSLIERYTEIDPSIEITICRGYKGLLGDSPVTAEVK 63
 QY 93 AGVLRFPGSGVIGNSRVKLTNNKDCVKRGLVKEGSDPOKADOLVKQVDILHTIGDD 152
 DB 64 AAVLHFGSGSPGNSRVKLTNNKDCVKRGLVKEGSDPOKADOLVKQVDILHTIGDD 123
 QY 153 TTTAAADLAFLARNNGYLVIGLPTVDNDVFPFKQSLGAWTAAEGARFPMVNAEN 212
 DB 124 TTTAAADLAFLARNNGYLVIGLPTVDNDVFPFKQSLGAWTAAEGARFPMVNAEN 183
 QY 213 ANPRMLTVEHWGRNCGMVLTAAADRYKRLDRAWMLBELGTRSEYHAFVPEMAID 272
 DB 184 AAVRLTVEHWGRNCGMVLTAAADRYKRLDRAWMLBELGTRSEYHAFVPEMAID 243
 QY 273 LEAEAKRLREVNDKVDVNIFFVSEGAGYEAIYAEMQAGOEVRDAFGHIKLDVNPGR 332
 DB 244 LEAEAKRLREVNDKVDVNIFFVSEGAGYEAIYAEMQAGOEVRDAFGHIKLDVNPGR 303
 QY 333 FGEQFQMIAGAEKTLVQKSGYFARASNVDDMRILKSCADLAVCAFRRESGVIGHED 392
 DB 304 FSKQFALLGABRSVQKSGYFARASNVDDMRILKSCADLAVCAFRRESGVIGHED 363
 QY 393 NGNVLRATLFPRIKGGKPPNITDTPFNMSLSEIQ 427
 DB 364 OGGRLLTLEFPRIKGGKPPNITDTPFNMSLSEIQ 398

RESULT 6
 Q8UDL0 PRELIMINARY; PRT; 479 AA.
 AC Q8UDL0;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Pyrophosphate--fructose-6-phosphate 1-phosphotransferase.
 GN Name=pfp; OrderedLocNames=Atu2115;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NCBI_TaxID=176299;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Dupont;
 RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavyn T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Seemphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Neeser E.W.,
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323 (2001).
 DR EMBL; AE009161; AAL4105.1; -.
 DR PIR; AC2836; AC2836.
 DR PIR; F97613; F97613.
 DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.

DR Pfam; PF00365; PRK; 1.
 DR PRINTS; PR00476; PHPRCKINASE.
 DR ProDom; PD000707; Pfruckinase; 1.
 KW Complete proteome; Transferase.
 SQ SEQUENCE 479 AA; 52036 MW; 86934EBF948A2C56 CRC64;
 Query Match 56.6%; Score 1292.5; DB 2; Length 479;
 Best Local Similarity 57.9%; Pred. No. 3.9e-73;
 Matches 246; Conservative 71; Mismatches 101; Indels 7; Gaps 2;

QY 8 HLTAADI-----RFGWPFLENFNYTLNNKPKVAILTAGLAPCLNSAIGSLIERYTEIDP 63
 DB 54 HLHAAVNSICOTCLMVAK---RRAMAKQKVAMLTAGLAPCLNSAIGSLIERSDIAP 110
 QY 64 SIEITICRGYKGLLGDSPVTAEVKKAQVLFQSGSVIGNSRVKLTNNKDCVKRGLV 123
 DB 111 EIDIIAYSSGYQGVILGRIETITDMREKAHLHRYGSSPIGNSRVKLTNNADCAKGLV 170
 QY 124 KEGSDPOKADOLVKQVDILHTIGDDTNTAAADLAFLARNNGYLVIGLPTVDND 183
 DB 171 KEGSDPPLVAERLAADGITLHTIGDDTNTAAADLAFLARNNGYLVIGLPTVDND 230
 QY 184 VFPIKQSLGAWTAAEGARFPMVNAENANPRMLTVEHWGRNCGMVLTAAADRYKRL 243
 DB 231 VVPIKQSLGAWTAAEGARFPMVNAENANPRMLTVEHWGRNCGMVLTAAADRYKRL 290
 QY 244 DRAEWLPELGLTRSEYHAFVPEMAIDLEAEAKRLREVNDKVDVNIFFVSEGAGYEAI 303
 DB 291 RGNVYVEELMNTMKNIDGITLPEMAFDIEAEAKRLREVNDKVDVNIFFVSEGAGYEAI 350
 QY 304 VAEMQAGOEVRDAFGHIKLDVNPGRKFGEQFQMIAGAEKTLVQKSGYFARASNVDD 363
 DB 351 VAERBAEAVKRAFGHVKIDITINVGWFOQKQFAGLGAERSVQKSGYFARASNPAGD 410
 QY 364 DMRILKSCADLAVCAFRRESGVIGHEDNGNVLRATLFPRIKGGKPPNITDTPFNMSLS 423
 DB 411 DLRIQGVNDLAVSALNKVSGVIGHEDNGKRLRTIEFPRIKGGKPPNITDTPFNMSLS 470

RESULT 7
 Q7CXV0 PRELIMINARY; PRT; 514 AA.
 AC Q7CXV0;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE AGR_C_3836P.
 GN OrderedLocNames=AGR_C_3836;
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 NCBI_TaxID=176299;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Cereon;
 RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Guttorlo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
 RA Houmlet K., Gordon J., Vaubin M., Iatchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328 (2001).
 DR EMBL; AE008127; AAK87863.1; -.
 DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.


```

Db 200 EQVCSWTSRFRASVAPI-----VVVAEGA-----MPROGDWVLKDESLDSYGHVRLS 247
QY 336 AVNPEKMFGEQPAQMTGAB-KTLVQKSGYFAPASASNVDDMKLILSCADLAVECAFRRBS 384
Db 248 GV--GEMLAKQIEKKTGNEARTTV--LGHVQCGTSPSAFDRWLARFGLHVAIDCV----- 298
QY 385 GVIGHDEDNQNV--LRATIEFPRI 405
Db 299 ---HGDGFGKVALRGTDIVRV 317

RESULT 9
PFP_DICTH STANDARD; PRT; 346 AA.
AC Q9K71,
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pyrophosphate--fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90)
DE (6-phosphofructokinase, pyrophosphate dependent) (pyrophosphate-
DE dependent 6-phosphofructose-1-kinase) (PPI-dependent
DE phosphofructokinase) (PPI-PFK).
GN Name=pfp;
OS Dicyoglomus thermophilum.
OC Bacteria; Dicyoglom; Dicyoglomales; Dicyoglomaceae; Dicyoglomus.
OX NCBI_TaxId=14;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rt46 B.1;
RX MEDLINE=20372656; PubMed=10913106;
RX DOI=10.1128/JB.182.16.4661-4666.2000;
RA Ding Y.-H.R., Kontius R.S., Morgan H.W.;
RT "Sequencing, cloning, and high-level expression of the pfp gene,
RT encoding a ppi-dependent phosphofructokinase from the extremely
RT thermophilic eubacterium Dicyoglomus thermophilum.";
RL J. Bacteriol. 182:4661-4666(2000).
CC -I- CATALYTIC ACTIVITY: Diphosphate + D-fructose 6-phosphate =
CC phosphate + D-fructose 1,6-bisphosphate.
CC -I- PATHWAY: Key control step of glycolysis.
CC -I- SIMILARITY: Belongs to the phosphofructokinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF268276; AAF80100.1; -.
DR HSSP; P06998; 2PFK.
DR HAMAP; MF_00339; -; 1.
DR InterPro; IPR000023; Ppfruckinase.
DR Pfam; PF00365; PFK; 1.
DR PRINTS; PR00476; PPHRCTKINASE.
DR PRODOM; PD000707; Ppfruckinase; 1.
DR DR ALOsteric enzyme; ATP-binding; Glycolysis; Kinase; Magnesium;
DR transferase.
KM NP_BIND 23
KM NP_BIND 156
FT NP_BIND 173
FT ACT_SITE 129
FT BINDING 164
FT BINDING 269
FT BINDING 275
FT BINDING 278
SEQUENCE 346 AA; 37448 MW; 50C03B64RA7927FI CRC64;
Query Match 13.4%; Score 305; DB 1; Length 346;
Beet Local Similarity 27.5%; Pred. No. 3e-11;
Matches 107; Conservative 73; Mismatches 155; Indels 54; Gaps 15;

```

```

QY 34 KVALITAGGLAPCLNSAIGSLIERYTEIDPSIRIIICVRGKGLLGDSPYTAVERKKA 93
Db 5 RIGVLITGGDCPGALPALRGIVMR--ADVDYDEYGLKRYGAGLLKADIMPLSLEWED- 61
QY 94 GVLORFGGSSVIGNSRVXLTVNKDCVKGLYKEDGEPQKRAADQVLKQDVLHTIGSDT 153
Db 62 --ILEIGTILGSSR-----TNP-----FKKEEDYQK-CVENPKKLLDLIALIGSDT 107
QY 154 NTAADLAFLARNNYGLTVIGLPRYDNDVFPPIKQSIGANTAAEQCARFEMNVVAENNA 213
Db 108 LGVASKF-----SKLGIPMIGVPTIKDKLEEDTYTGPTAVEVVDAIKRLDARS 161
QY 214 NPMILVHEVMGRNCGMTAATAOEYRKLDRLEMLPELGLTRSEYVAHAFVPEMADL 273
Db 162 HARVIVV-ELMGRIAGLAL-----YGLAGADY-----ILIEVERNL 200
QY 274 EAEARKLEAVDK-VDCNVIFVSEAGVEALVAEQAQGVPRDPAFGIKLDAVNPGRK 332
Db 201 EDLVNHRKLVARGRNNAVVAIAEGVQLPGFTYQ---KGQGMVDAFCHIRLGGV--GNV 255
QY 333 FGEQPAQMTGAEKTLVQKSGYFAPASASNVDDMKLILSCADLAVECARRESGVGHED 392
Db 256 LAEBIQKVLGIETRAVLIS-HLQKGGSPSIRDRIMGILLGKKAVDLVHEKSGLF--VAV 312
QY 393 NGNVLRATIEFPRIKQKPEFIDTDMFNSM 421
Db 313 KGNELVPVDTILIE-GKTRVNDPAFYESV 340

RESULT 10
PFP_AMEYE STANDARD; PRT; 341 AA.
AC Q59126,
DT 15-JUL-1998 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pyrophosphate--fructose 6-phosphate 1-phosphotransferase (EC 2.7.1.90)
DE (6-phosphofructokinase, pyrophosphate dependent) (pyrophosphate-
DE dependent 6-phosphofructose-1-kinase) (PPI-dependent
DE phosphofructokinase) (PPI-PFK).
GN Name=pfp;
OS Amycolatopsis methanolica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
OX NCBI_TaxId=1814;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96125240; PubMed=8550409;
RA Alves A.M., Meijer W.G., Vrijbloed J.W., Dijkhuizen L.;
RT "Characterization and phylogeny of the pfp gene of Amycolatopsis
RT methanolica encoding ppi-dependent phosphofructokinase.";
RL J. Bacteriol. 178:149-155(1996).
CC -I- CATALYTIC ACTIVITY: Diphosphate + D-fructose 6-phosphate =
CC phosphate + D-fructose 1,6-bisphosphate.
CC -I- PATHWAY: Key control step of glycolysis.
CC -I- SUBUNIT: Homotrimer.
CC -I- SIMILARITY: Belongs to the phosphofructokinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U31277; AAB01683.1; ALT_INIT.
DR HSSP; P00512; 3PFK.
DR HAMAP; MF_00339; -; 1.
DR InterPro; IPR000023; Ppfruckinase.
DR Pfam; PF00365; PFK; 1.
DR PRINTS; PR00476; PPHRCTKINASE.
DR PRODOM; PD000707; Ppfruckinase; 1.

```


RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.
 RT "Genome sequence of an industrial microorganism *Streptomyces*
 RT *avermitilis*: deducing the ability of producing secondary
 RT metabolites.";
 RA Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WA-4680;
 RX MEDLINE=22608306; Pubmed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism *Streptomyces avermitilis*.";
 RL Nat. Biotechnol. 21:526-531(2003).
 CC -1- PATHWAY: Key control step of glycolysis.
 CC -1- SIMILARITY: belongs to the phosphofructokinase family.
 DR EMBL; AP005049; BAC74834.1; -
 DR HSSP; P00512; 3PRK.
 DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO; GO:0003872; P:6-phosphofructokinase activity; IEA.
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000023; Pf:fructokinase.
 DR Pfam; PF00365; PRK; 1.
 DR PRINTS; PR00476; PHERCTKINASE.
 DR Prodom; PD00707; Pf:fructokinase; 1.
 DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 1.
 DR Complete proteome; Glycolysis; Kinase; Transferase.
 SQ SEQUENCE 341 AA; 36425 MW; 2B24F501DD862D6 CRC64;
 Query Match 12.5%; Score 284.5; DB 2; Length 341;
 Best Local Similarity 27.5%; Pred. No. 5,8e-10;
 Matches 97; Conservative 62; Mismatches 143; Indels 51; Gaps 11;
 QY 34 KVAITTAGGLACLSAIGSLIERTEIDPSIEIICYRGYKGLLGDSYPTAEVRKKA 93
 DB 2 RGVLTSGGDCGLNVRISVVR-AVVDHGEVIGFRDGMGLGECDD--LKLDVAVS 58
 QY 94 GVLQFFGSGVINSRVKLTNVDCVKGGLVKEGDPQKVAADQLVNDGVDILHTIGDDT 153
 DB 59 GILAR-GGITLSSSRVQPAHLRDSVER-----ARGHVALGLDAIPIIGSEGT 105
 QY 154 NTAADLAFAFLARNYGLTVIGLPTVDNDVPIKOSLGAMTAAGGARFVNVVAENNA 213
 DB 106 LKAAALL-----SDAGLPIVGVPTKIDINDIATVDTFGFDPAVGAVTEALDKTTAES 159
 QY 214 NFRMLIVHEVMGRCGMLTAATAOEVKRLDRAEMLPGLTRESVEYVAVPEMAIDL 273
 DB 160 HORVLLV-EVMGRHTGMIALHSGM-----AAGAHAIIVPEPFDI 198
 QY 274 EAAARLREVMKVDVCNIFV-SEGAGVEAIVAEQAQGVPRPAFGHKLDAVNPGRK 332
 DB 199 BELAAKVGSRFEAGKRFALIVAAEGAKPRAGSMDF-----DEGKKDYGHERRAGI--ARQ 252
 QY 333 FEEQPAQMIGAETLVQKSGYFAPASASVNDMLIKSCADIAVECAFRESG 385
 DB 253 LSLLEERLIGKEARPV-ILGHVQGGTPTAYRVLATRFGMAHVAEVAHGEFG 304
 RESULT 13
 Q9FJUS PRELIMINARY; PRT; 488 AA.
 ID Q9FJUS
 AC Q9FJUS
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
 DE Pyrophosphate-dependent phosphofructo-1-kinase-like protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosidII; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98403884; Pubmed=9734815;
 RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneo T., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:203-216(1998).
 DR EMBL; AB013392; BAB09881.1; -
 DR HSSP; P00512; 3PRK.
 DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO; GO:0003872; P:6-phosphofructokinase activity; IEA.
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000023; Pf:fructokinase.
 DR Pfam; PF00365; PRK; 1.
 DR PRINTS; PR00476; PHERCTKINASE.
 DR Prodom; PD00707; Pf:fructokinase; 1.
 KW Kinase.
 SQ SEQUENCE 488 AA; 53781 MW; ABA526AED923B17 CRC64;
 Query Match 12.3%; Score 282; DB 2; Length 488;
 Best Local Similarity 26.8%; Pred. No. 1.3e-09;
 Matches 112; Conservative 60; Mismatches 158; Indels 88; Gaps 18;
 QY 37 ILTAGGLAPCLNSAIGSL-----IERYTEIDPSIEIICYRGYKGLLGDSYPTA 87
 DB 96 IYTCGGLCPGLNTVIREVSSLSVWYGVKRLIGDIVSL-----CGYGFYAKNTIPLNS 149
 QY 88 EVRKAGVQLQFFGSGVINSRVKLTNVDCVKGGLVKEGDPQKVAADQLVNDGVDILHT 147
 DB 150 KV---VNDIHRGGTITGTSR-----GSHDTNKI-VSDIDRGINQYIT 189
 QY 148 IGGDDTNTAAADLAFAFLARNYGLTVIGLPTVDNDVPIKOSLGAMTAAGGARF--M 205
 DB 190 IGGDGTGAGAVIPEIRRLRLKAVVGIPKTIIDVIDVSKSFGFDPAVEAQAQAINNA 249
 QY 206 NVVAENNAFPMILIVHEVMGRCGMLTAATAOEVKRLDRAEMLPGLTRESVEYVAHV 265
 DB 250 HVEASNNENGIGFV--KLMDGYSGVIA-----MYATLASRVDVDDL 288
 QY 266 VPMAIDLEAA-----KLRBYMDKYDCNIFVSEGAGVEAIVAEQAQGVPRPA 318
 DB 289 IPESEFYLEGEGGLFEFIERRLKDHGVN---IVLAGAGODMLCKSNES---TPMDA 340
 QY 319 FGHILDAVNPCKMFGSEQPAQWIGAETLVQ-----KSGYFAPASASVNDMLIKSCADL 374
 DB 341 SGNKLKRV--GLMISQSIKHFKKRVNMLKTYDPTMYIRAVSNASNDVYCTLLAQS 398
 QY 375 AVECAFRESGVIGHED--NGNVLRALFPPRIKGGKPFNIDTD--WFNSMLSEIGOP 428
 DB 399 AVHGAM---AGYGTYSGLVNGR-QTYIPFYRITETQNNVVITDMMAR-LLSTNOP 451
 RESULT 14
 K6P3 STRCO STANDARD; PRT; 341 AA.
 ID K6P3 STRCO
 AC Q9FC39;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE 6-phosphofructokinase 3 (EC 2.7.1.11) (Phosphofructokinase 3)
 DE (Phosphohexokinase 3).
 GN Name=pfk3; Synonyms=pfk3; OrderedLocusNames=SCO1214;
 GN ORNames=25C058.14;
 GN Streptomyces coelicolor.
 OC Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OK NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;

MEDLINE:21996410; PubMed:12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Gerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleefer H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RA "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)";
 RL Nature 417:141-147(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
 CC fructose 1,6-bisphosphate.
 CC -1- PATHWAY: Key control step of glycolysis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the phosphofructokinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF39108; CAC01496.1; -.
 DR HSSP: P00512; 3PFK. -.
 DR HAMAP: MF_00339; 1.
 DR InterPro: IPR000023; Ppfuckkinase.
 DR Pfam: PF00365; PFK, 1.
 DR PRINTS: PR00476; PPFCKTKINASE.
 DR ProDom: PD000707; Ppfuckkinase. 1.
 DR PROSITE: PS00433; PHOSPHOFRUCTOKINASE. 1.
 DR Allsteric enzyme; ATP-binding; Complete proteome; Glycolysis; Kinase;
 KM Magnesium; Multigene family; Transerase.
 FT NP_BIND 20 24 ATP (By similarity).
 FT NP_BIND 154 158 ATP (By similarity).
 FT NP_BIND 171 187 ATP (By similarity).
 FT ACT_SITE 127 127 Proton acceptor (By similarity).
 FT BINDING 162 162 Substrate (By similarity).
 FT BINDING 266 266 Substrate (By similarity).
 FT BINDING 272 272 Substrate (By similarity).
 FT BINDING 275 275 Substrate (By similarity).
 FT SEQUENCE 341 AA; 36431 MW; E3050D37BDB6P9F0 CRC64;
 SQ
 Query Match 12.3%; Score 281.5; DB 1; Length 341;
 Best Local Similarity 26.8%; Pred. No. 9e-10;
 Matches 100; Conservative 66; Mismatches 152; Indels 55; Gaps 12;
 QY 34 KVALTLAGLAPLNSAIGSLERYTEIDPSEIILCYRGYGLLGDSVPYTAERKKA 93
 DB 2 RIGVLISGGDCPELNAINRSVVMR-AVVDGDEVIFRGMKGLTCDY--LKLDDAVG 58
 QY 94 GVLQREGSGVIGNSRYKLTNVKDCVKGEGEDPQKVAADQLVNDGVYDLHTTGSDT 153
 DB 59 GLIAR-CGIIILSSSRPRPHLRGVER-----ARGHVELGLDIIITPGEGT 105
 QY 154 NTPAADLAAFLARNNYGLTVIGLPTVDNDVFPFKQSGLGAMTAEGGARYFNNVNAENNA 213
 DB 106 LKAAARLL-----SDNGLPIVGPPTIIDNDIATDVTGFDPTAVATBALDLKTTAAS 159
 QY 214 NPMRLIVHEVMGRNCMLTAATQAEYRKLLDRAEWMLPELGLTRESVEYNAVPEMAIDL 273
 DB 160 HQRVLV-LVEMGHMTWIALHSGM-----AAGAAVAVVPERPFI 198
 QY 274 EAAKALREVMDKVCNIFV-SEGAGVEAIVAEQAKQOEVRDFAFGHKLDAVVPKKN 332
 DB 199 DELTAVGRFSAKGRFALVVAEGAKPKAGTMDPDEGG---KDYVGHERRFAGI--AQ 252
 QY 333 FEGQFQMTGAETKTVQKSGYFAPASASVNDMLIKSCADLAVECAFRESGVI---G 388

DB 253 LSEIEERLKEKAPRV-ILGHVQGGTPTAYDRVLATRFGMHVAEVAHERGEGKOTALNG 311
 QY 389 HDENGNVRATRE 401
 DB 312 TLIEWSLADAVE 324
 RESULT 15
 ID 09CSJ7 PRELIMINARY; PRT; 485 AA.
 AC 09CSJ7;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Putative pyrophosphate-dependent phosphofructo-1-kinase.
 GN Name=At5g5630;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosoid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 RA Carlinici P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Egu P., Lee J.M.,
 RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
 RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shim P.,
 RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
 RL EMBL: AF360207; MAK25917.1; -.
 DR EMBL: AY040055; AAK64113.1; -.
 DR HSSP: P00512; 1MTO.
 DR GO: GO:0005945; C:6-phosphofructokinase complex; IEA.
 DR GO: GO:000872; F:6-phosphofructokinase activity; IEA.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:006096; P:glycolysis; IEA.
 DR InterPro: IPR000023; Ppfuckkinase.
 DR Pfam: PF00365; PFK, 1.
 DR PRINTS: PR00476; PPFCKTKINASE.
 DR ProDom: PD000707; Ppfuckkinase. 1.
 KM Kinase.
 SQ SEQUENCE 485 AA; 53482 MW; 6F0C6CF93EAA7B5 CRC64;
 QY 37 IITAGSLAPLNSAIGSLERYTEIDPSEIILCYRGYGLLGDSVPYTAERKKA 96
 DB 96 IYTCGLCPGLNTVIREVSSLSYMGVKKIIGDGGIRGFPAKTIPLNSKV--VNDI 152
 QY 97 QRFSGSVIGNSRYKLTNVKDCVKGEGEDPQKVAADQLVNDGVYDLHTTGSDTNTA 156
 DB 153 HKRGITIGTSR-----CGHDTNNTI-VDSIDRGINQYIITGGDTG 195
 QY 157 AADLAAFLARNNYGLTVIGLPTVDNDVFPFKQSGLGAMTAEGGARYF--NNVNAENNA 214
 DB 196 ASVIFEEIRRRRLKAAVVGIPPTIINDIPVIDKSGFPTAVEAQRATNAHAHVEAESNEN 255
 QY 215 PMRLIVHEVMGRNCMLTAATQAEYRKLLDRAEWMLPELGLTRESVEYNAVPEMAIDL 274
 DB 256 GIGFV--KUMGRYSGITA-----MATLASDVOVCCLLPESPFYIE 294
 QY 275 AEA-----KRLREVMDKVCNIFVSEGAGVEAIVAEQAKQOEVRDFAFGHKLDAV 327

```

Db      295 GEGGLPEPTERRLKDHGHNV---IVLAAGAGODLMCKSMES---TPMDASGNKLLKDV 346
Oy      328 NPGKNFGEOFAQMIGAETLVQ---KSGYFARASASNVDDRLIKSCADLAVECAPRRE 383
Db      347 --GLMSQSIXDHFKNKVMMLKYIDPTYMIRAVPSNASDNYCTLLAQSAVHGAM--- 401
Oy      384 SGVIGHDED--NGNVLRAIEPPRIKGGKPFENIDTD--WNSMLSEIGQP 428
Db      402 AGYGTSGLVNGR-QTYIPFYRITETQNNVVITDRMMAR-LLSSTNQF 448

```

Search completed: February 24, 2005, 14:22:28
 Job time : 185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 14:01:54 ; Search time 40 Seconds

(without alignments)
1051.168 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2264
Sequence: 1 DVVTWPHYLTDIRFCWFF.....FNSMSEIQPKGKRVESH 437

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|---------------------|
| 1 | 1314 | 57.5 | 405 | 2 A41169 | diphosphate-fructo |
| 2 | 1292.5 | 56.6 | 479 | 2 AC2836 | hydrophatical prote |
| 3 | 1292.5 | 56.6 | 514 | 2 P97613 | ppl-phosphofructok |
| 4 | 311 | 13.6 | 342 | 2 T35500 | 6-phosphofructokin |
| 5 | 270 | 11.8 | 448 | 1 F70190 | probable diphospha |
| 6 | 263.5 | 11.5 | 436 | 1 S49458 | diphosphate-fructo |
| 7 | 257.5 | 11.3 | 500 | 2 T06011 | probable diphospha |
| 8 | 253 | 11.1 | 319 | 2 C72406 | 6-phosphofructokin |
| 9 | 251 | 11.0 | 473 | 2 T13433 | pyrophosphate-depe |
| 10 | 246.5 | 10.8 | 987 | 1 UQ0016 | 6-phosphofructokin |
| 11 | 245.5 | 10.7 | 992 | 2 S32902 | 6-phosphofructokin |
| 12 | 243 | 10.6 | 462 | 2 T10691 | probable diphospha |
| 13 | 241.5 | 10.6 | 437 | 2 S54978 | 6-phosphofructokin |
| 14 | 240.5 | 10.5 | 320 | 2 AH0942 | 6-phosphofructokin |
| 15 | 238.5 | 10.4 | 320 | 1 KIECPA | 6-phosphofructokin |
| 16 | 238.5 | 10.4 | 320 | 2 A86081 | 6-phosphofructokin |
| 17 | 238.5 | 10.4 | 320 | 2 A98234 | 6-phosphofructokin |
| 18 | 234.5 | 10.3 | 327 | 2 AF0010 | 6-phosphofructokin |
| 19 | 233.5 | 10.2 | 322 | 2 S39248 | 6-phosphofructokin |
| 20 | 233 | 10.2 | 319 | 2 S35928 | 6-phosphofructokin |
| 21 | 232 | 10.2 | 461 | 1 A71366 | diphosphate-fructo |
| 22 | 226 | 9.9 | 322 | 1 UQ1028 | 6-phosphofructokin |
| 23 | 224.5 | 9.8 | 334 | 2 A75495 | 6-phosphofructokin |
| 24 | 223.5 | 9.8 | 319 | 2 D84045 | 6-phosphofructokin |
| 25 | 223.5 | 9.8 | 331 | 2 G82045 | 6-phosphofructokin |
| 26 | 223 | 9.8 | 343 | 2 E70856 | probable pfka prot |
| 27 | 222 | 9.7 | 340 | 1 JN0614 | 6-phosphofructokin |
| 28 | 221 | 9.7 | 335 | 2 D97971 | 6-phosphofructokin |
| 29 | 220 | 9.6 | 573 | 2 C71312 | probable pyrophosp |

| | | | | | |
|----|-------|-----|-----|----------|--------------------|
| 30 | 220 | 9.6 | 780 | 1 KIRBF | 6-phosphofructokin |
| 31 | 216.5 | 9.5 | 319 | 2 P96963 | 6-phosphofructokin |
| 32 | 215 | 9.4 | 335 | 2 P95103 | 6-phosphofructokin |
| 33 | 213.5 | 9.3 | 321 | 2 C70447 | phosphofructokinas |
| 34 | 211.5 | 9.3 | 348 | 2 T45407 | phosphofructokinas |
| 35 | 210.5 | 9.2 | 320 | 2 P84965 | 6-phosphofructokin |
| 36 | 210.5 | 9.2 | 343 | 2 G87121 | 6-phosphofructokin |
| 37 | 210.5 | 9.2 | 544 | 2 S52081 | diphosphate-fructo |
| 38 | 210 | 9.2 | 780 | 1 KIRUFM | 6-phosphofructokin |
| 39 | 210 | 9.2 | 789 | 1 A45617 | 6-phosphofructokin |
| 40 | 209 | 9.2 | 319 | 1 KIRSF | 6-phosphofructokin |
| 41 | 209 | 9.2 | 323 | 1 G64223 | 6-phosphofructokin |
| 42 | 207.5 | 9.1 | 791 | 2 A53206 | 6-phosphofructokin |
| 43 | 207 | 9.1 | 959 | 1 JQ0017 | 6-phosphofructokin |
| 44 | 204.5 | 9.0 | 784 | 2 JC2055 | 6-phosphofructokin |
| 45 | 204 | 8.9 | 463 | 2 B84613 | hypothetical prote |

ALIGNMENTS

RESULT 1

A41169
diphosphate-fructose-6-phosphate-1-phosphotransferase (EC 2.7.1.90) - Propionibacterium f
N:Alternate names: 6-phosphofructokinase (pyrophosphate)
C/Species: Propionibacterium freudenreichii
C/Date: 17-Jul-1992 #sequence_revision 15-Mar-2004 #ext_change 09-Jul-2004
C/Accession: A41169
R/radior, U.S.; Gollapudi, L.; Tripathi, R.L.; Latehaw, S.P.; Kemp, R.G.
U. Biol. Chem. 266, 16550-16555, 1991
A/Title: Cloning, sequencing, and expression of pyrophosphate-dependent phosphofructokin
A/Reference number: A41169; PMID:9158443; PMID:1653240
A/Molecule type: DNA
A/Residues: 1-404 <LAD>
A/Cross-references: UNIPROT:P29495; GB:M67447; NID:g150930; PIND:AAA2675.1; PID:g150931
C/Comment: This is an example of a nonallosteric, pyrophosphate-dependent phosphofructokir
C/Superfamily: pyrophosphate-dependent phosphofructokinase, SMC01852 type
C/Keywords: phosphotransferase
F/5-322/Domain: 6-phosphofructokinase 1 homology <6PPI>

Query Match 57.5%; Score 1314; DB 2; Length 405;
Best Local Similarity 62.6%; Pred. No. 7, 8e-85;
Matches 248; Conservative 60; Mismatches 88; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 33 | KVAVITAGGLAPCNSAIGSLIERYTEIDPSLEIICVGGYKGLIGDSYPTAEVRKK | 92 |
| DB | 4 | KVAVITAGGLAPCNSAIGSLIERYTEIDPSLEIICVGGYKGLIGDSYPTAEVRKK | 63 |
| QY | 93 | AGVLRFGSVIGNSRVKLTNVKDCVKGVLVEGEDPQKVAADQVLKDGVDILHTIGGD | 152 |
| DB | 64 | YRLRFGSGSPIGNSVKLTNVKDLVARKLVASGDPLKVAADQVLADVDVHTIGGD | 123 |
| QY | 153 | TVTAAADLAFLARNNGYLVITGLPPTVDNDVFPKQSIGAVTAEGARFENVVAENN | 212 |
| DB | 124 | TVTAAADLAFLARNNGYLVITGLPPTVDNDVFPKQSIGAVTAEGARFENVVAENN | 183 |
| QY | 213 | ANPRMLIVHEWNGRNGMLTAATAQERYKGLDRAEKLPELGLTREYVHAHFVPEMAD | 272 |
| DB | 184 | ANPRELIHEIRNGRCYGLAATSRRYVAMLAQOGLPAGADRGMIDHILVYPEATID | 243 |
| QY | 273 | LEAEAKLEBNDKVDGVNI FVSEGAGVEAIVAEOMAKGOEYPRDAFGIKLDAVPGKW | 332 |
| DB | 244 | LDABERAKLTVDGVSVIIFISEGAGVDIYAQMOKATQEPPTDAFGVQDKINPGAW | 303 |
| QY | 333 | FEQPAQMIAGKTLVQKSGYFAPASASVNDMDRLIKSCADLAIVECAFRRSGVIGHSD | 392 |
| DB | 304 | FAKQFAERIGAGKTVQKSGYFSRSKVAQDLLELAATVATVADLAAGTGGVGDDBE | 363 |
| QY | 393 | NGNVLRATFRPRIKGKPEINIDTDFNSMLSTICGP | 428 |
| DB | 364 | ADDKLSVIDFKRIAGHKPPDITLDWYTQLARIQGP | 399 |

AC2836
 AC2836
 hypotheoretical protein pfp (imported) - Agrobacterium tumefaciens (strain C58, Dupont)
 C/Species: Agrobacterium tumefaciens
 C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C/Accession: AC2836
 R/Mood, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McEllell
 Science 294, 2317-2323, 2001
 A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A/Reference number: AB2577; MUID:21608550; PMID:11743193
 A/Accession: AC2836
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-479 <KIR>
 A/Cross-references: UNIPROT:Q8UDL0; GB:AB008688; PIDN:AAL43105.1; PID:G17740576; GSPDB:G
 C/Genetic source: strain C58 (Dupont)
 A/Genetic source: strain C58 (Dupont)
 A/Map position: circular chromosome
 C/Superfamily: pyrophosphate-dependent phosphofructokinase, SMC01852 type; 6-phosphofruc
 Query Match 56.6%; Score 1292.5; DB 2; Length 479;
 Best Local Similarity 57.9%; Pred. No. 3.2e-83;
 Matches 246; Conservative 71; Mismatches 101; Indels 7; Gaps 2;
 QY 8 HL7ADI---RFGHWFLENFNYTLNKKPKKVALITLGGIAPCLNSAIGSLIERYTEIDP 63
 DB 54 HLHAAVASISQCTCLMVYAK---RRAMMAKQKVALITLGGIAPCLNSAIGSLIERYSIDAP 110
 QY 64 SLEITICRGVKKGLLDSDSYPTAEVRKKAGVLPFGSGVYGNRKYLTANKDCVRGCV 123
 DB 111 EIDIIAARSGYQGVLLGERIEITKDMREKALHLRHYGSPYGNRKYLTNAADAKRGGLV 170
 QY 124 KEGEDPOKVAADQVVKDGVLDLHTIGGDDTTAAADLAFLARNNGYLVGLPKTYDND 183
 DB 171 KEGNNPLRVAERLAADGTTILHTIGGDDTTAAADLAAYIGANGYDVLVGLPKTYDND 230
 QY 184 VPIKOSLGAAVTAEGARVFNVAENNANRMLIYHEVGNRCGMVLTAAQGYRKL 243
 DB 231 VVPIKOSLGAAVTAEGARVFNVAENNANRMLIYHEVGNRCGMVLTAAQGYRKL 290
 QY 244 DRAMWELBELGTRSYEVHAFVPEMAIDLEAEKRLREVDKDCVNIPISEAGYEAI 303
 DB 291 RGNVYEGGLMMNTOMKINDGILPEEMAFDIEAEERLKEVMDKRGYVTLFVSEAGYDAI 350
 QY 304 VAEQDAQGOEVPDAFGHIIKLDVAVPCKFGEQFQOMIGAEXTLVOKSGVFARASAGNV 363
 DB 351 VAEEAAGAEAKRPAFGHVKIDITINVGWPKQKAGLIGARSNVQKSGVFARASAGNV 410
 QY 364 DMRLIKSCADLAVECAFRRESGVYGHDENGNVLRALIEFPRIKGGKPPNIDTDFNSMLS 423
 DB 411 DLRLIQGVNDAVSAALNKVSGVYGHDENGNVLRALIEFPRIKGGKPPNIDTDFNSMLS 470
 QY 424 EIGOP 428
 DB 471 HVGOP 475
 RESULT 3
 F97613
 ppi-phosphofructokinase (AF246209) (imported) - Agrobacterium tumefaciens (strain C58, C
 C/Species: Agrobacterium tumefaciens
 C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C/Accession: F97613
 R/Goodner, B.; Hinkley, G.; Galtung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A975359; MUID:21608551; PMID:11743194
A:Accession: P97613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-514 <KUR>
A:Cross-references: UNIPROT:O8UDL0; GB:AEO07869; PIDN:AAK87863.1; PID:g15157249; GSPDB:GN
C:Genetics:
A:Gene: AGR_C_3836
A:Map position: circular chromosome
C:Superfamily: pyrophosphate-dependent phosphofructokinase, SMC01852 type

Query Match 56.6%; Score 1292.5; DB 2; Length 514;
Best Local Similarity 57.9%; Pred. No. 3.5e-83;
Matches 246; Conservative 71; Mismatches 101; Indels 7; Gaps 2;

OY HLTADI---RFGHWFLENFNYTLANKPKVAILTAGGLAPCLNAISGLIRRYEIDP 63
Db HLHAHAVASICOTCTLMVVAK---RRAMMAKQKVAMLTAGLAICTLSAVGGLIERYSDIAP 145
64 SEIILCYRGKGKGLIGDSYPVTAEVRKKAGVLORPGSGVINSRYKLITNVKOCVRGLV 123
OY EIDIILAIRSGYGQGVLLGERIEITTKDMREKAMHLHYGGSPIGNSRYKLITNAADCAKGLV 205
Db EIDIILAIRSGYGQGVLLGERIEITTKDMREKAMHLHYGGSPIGNSRYKLITNAADCAKGLV 205
124 KEGEDPOKVAADOVKDQVDILHTIGDDDTNTAADLAAFARNNGYLVIGLPKTVDND 183
OY KEGDNPLRVAAERLAADDITLIHTIGDDDTNTAADLAAVLGANGDYLVGGLPKTYDND 265
Db KEGDNPLRVAAERLAADDITLIHTIGDDDTNTAADLAAVLGANGDYLVGGLPKTYDND 265
184 VPIIKOSIGAWTAAEQGARVFNNVVAENNANPRMLIVHEWGRNCGLTPATIAOEYRKLL 243
Db VPIIKOSIGAWTAAEVGAAPFDNVSNESAAPKFTVIHVEVMGRHCGLTPATIARAYIOQT 325
244 DRAEWLPELGITRESYEVAHFVPEMAYDLDEAEARLRPVMDVCVNIFVSAGAVEAI 303
Db RGDVVEELMNNTQKNIDGYLLPEMAPFIDEAEARLRKVPMDHGYYTVLFVSEAGIDLAI 385
OY VAEMQAKGOEVRPDAFGHIKILDVAVPGKMFGEQFMIGAEKTLVOKSGYFAPASANYD 363
Db VAEMQAKGOEVRPDAFGHIKILDVAVPGKMFGEQFMIGAEKTLVOKSGYFAPASANYD 363
386 VAEMQAKGOEVRPDAFGHIKILDVAVPGKMFGEQFMIGAEKTLVOKSGYFAPASANYD 445
OY DMRLIKSCADLAIVECAFRRREGVGIVGHDEDNGNLPAIEFPRIKGGKEPNITDWPMMSLS 423
Db DLRLIQGVNDLAVSALNAKVGSVTGHDDEQNKLTRIEFPRIKGGKHFLSAKMFEVME 505
OY 424 EIQGP 428
Db 506 HVGP 510

RESULT 4
T35500
6-phosphofructokinase (EC 2.7.1.11) - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C/Accession: T35500; T42063
J/Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21580
A:Accession: T35500
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <SES>
A:Cross-references: UNIPROT:O08333; EMBL:AL109661; PIDN:CAB51967.1; GSPDB:GN00070; SCODEB:
Appl. Environ. Microbiol. 63, 956-961, 1997
A>Title: Identification of ATP-dependent phosphofructokinase as a regulatory step in the
A:Accession: T42063
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <ALT>
A:Cross-references: EMBL:U51728; NID:g1931572; PIDN:AAC45135.1; PID:g1931573
C:Genetics:


```

QY 212 NANPMLIVHEVMGRCGLTAATAOERYKLLDRAEMLEBELGTRESYEVHAFVPEMAI 271
DB 242 KSAKRGIGIVRLMGDRDAGFIAL-----YASLANG-----DANLVLIPEIDI 282
QY 272 DL-----EAEAKRLREVMDVDCNIFVSEAGVEAIVAEOMAKGQEVPRDAFGHI-KLDA 326
DB 283 PRTQCEPFGK-----IMSGHYV-IVVAGA-----LQNKQKDDLDIGTSGNLIHDS 333
QY 327 VNPGRFGEQFAQMIGAETLVQ--KSGYFAPASASNVDDMLIKSCADLAVECAFRRRS 384
DB 334 INYLRDSITRYKLSIGIEBHITIKFVDPSPYMRSAFADAFHMCMLANAAVHVAMAGKT 393
QY 385 G-VIGHDENGVNLAIEFPR-----IKGKRPNIDTDV-FNSMLEIGOPK 429
DB 394 GLVICHNNHNN---FVSVPIRDTSTYIK-----RVMTDGLYTMASAIERPK 436

RESULT 7
T06011
Probable diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90) T25K17.80 -
N.Alternate names: protein T25K17.80
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
R.Bevan, M.; Rose, M.; Hempel, S.; Entlan, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, March 1999
A.Accession: T06011
A.Reference number: Z15382
A.Molecule type: DNA
A.Residues: 1-500 <BEV>
A.Cross-references: UNIPROT:Q9STQ7; EMBL:AL049171; GSPDB:GN00062; ATSP:T25K17.80
A.Experimental source: cultivar Columbia; BAC clone T25K17
C.Genetics:
A.Gene: ATSP:T25K17.80
A.Map position: 4
A.Introns: 46/2; 59/3; 80/3; 141/3; 194/3; 212/3; 235/3; 279/1; 367/3; 383/1; 434/3
C.Superfamily: pyrophosphate-dependent phosphofructokinase, TP0108 type; 6-phosphofructo
C.Keywords: phosphotransferase

Query Match 11.1%; Score 257.5; DB 2; Length 500;
Best Local Similarity 25.6%; Pred. No. 1.9e-10;
Matches 109; Conservative 59; Mismatches 167; Indels 91; Gaps 16;

QY 37 ILTAGLAPCLNSAIGSL-----IERYTEIDPSIEIIC-----YRGYKGL 78
DB 92 IVTCGGLCPGLMTVIREIVSSLSVMGYKRIIGIDVS--FCDCNLLLTKTNTGGRGYRFX 148
QY 79 LGDSYPTVAEVRKKAQVLRFGSGVIQNSRVLTIVKOCVKGKGLVKEGEDPQKVAADQLV 138
DB 149 AKNTVSLSSKV---VNDIHKRGSTILGTSR-----GGHDTTKI-VDSIQ 188
QY 139 KDGVDLHTIGDDNTTAADLAFLARNNYGLTVIGLPTVNDVFPFKOSLGAMTAAE 198
DB 189 DRGIVQVITIGDDGQGRGASVIFEEIRRGKLVANVIGIPKTIQNDIPIVDSFGFDVAE 248
QY 199 QGARYFMNVVANNANPMLIVHEVMGRCGLTAATAOERYKLLDRAEMLEBELGTRES 258
DB 249 EAQRIANNAHVAEBSIENIGIVKLMGRYSGFIA-----MYATLAS 289
QY 259 YEVAHVFPPEMAIDLEAE-----KRLREVMDVDCNIFVSEAGVEAIVAEOMAKG 311
DB 280 RDVDCCLIPESPFYJEGEGGLFEYIEKRLKESGHNV---LVIAAGAGDGLSKSMESHS- 344
QY 312 QEVPRDAFGHIKLDVNPGRFGEQFAQMIGAETLVQ--KSGYFAPASASNVDDMLRL 367
DB 345 --TLMDASGNKLLKQV--GLWISQSIKHFNOKKRVNMLKTYDPTMYIRAVSNASNDVY 400
QY 368 IKSQCADLAVECAFRRRSVIGHEDENGNY--LRAIEPRIKGKRPNIDTDV-FNSML 422
DB 401 CTLLAQSAVHGAMAGTYGI-----SGLVNGRQTYIPFYRIKONHNVITDRMMAR-LT 454
QY 423 SEIGOP 428

```

```

DB 455 SSTNQP 460

RESULT 8
C72406
6-phosphofructokinase - Thermotoga maritima (strain MSB)
C.Species: Thermotoga maritima
C.Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C.Accession: C72406
R.Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A.Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A.Reference number: A72200; MUID:99287316; PMID:10360571
A.Accession: C72406
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-319 <ARN>
A.Cross-references: UNIPROT:Q9WY52; GB:AE001705; GB:AE000512; NID:94980694; PIDN:AAJ35301
A.Experimental source: strain MSB
C.Genetics:
A.Gene: TM0209
C.Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase
F:4-279/Domain: 6-phosphofructokinase 1 homology <6FP>

Query Match 11.1%; Score 253; DB 2; Length 319;
Best Local Similarity 26.6%; Pred. No. 2.1e-10;
Matches 95; Conservative 60; Mismatches 126; Indels 76; Gaps 14;

QY 33 KKVAILTNGSLAPCLNSAIGSLIERYTEIDPSIEICRGYKGLLDSDSYPTAEVRKK 92
DB 2 KKIIVLTSGDAPGNNAVAV--RY-GVRQGLEVIQVRGSGSIDDPF--VKLEYKDV 57
QY 93 AGVLRFGSGVIQNSRVLTIVKOCVKGKGLVKEGEDPQKVAADQLVKQGVDLHTIGDD 152
DB 58 AGITRK-GGITLRSRCFEFT-----EEGELAAKQIKKGIIEGLVIGSEG 104
QY 213 ANPMLIVHEVMGRCGLTAATAOERYKLLDRAEMLEBELGTRESYEVHAFVPEMAID 272
DB 160 SHERAFIV-EVMGRHSGYIALMAG--LVTGAE-----AIYVEIPVD 198
QY 273 LEAEAKRLREV--MDKVCVNI FVSEAGVEAIVAEOMAKGQEVPRDAFGHIKLDVAVNG 330
DB 199 YSQLADRILIEERRKINSI-IVAEAGASAYTVAR-----HLE----- 236
QY 331 KMFGEQFAQMIGAETLVQKSGYFAPASASNVDDMLIKSCADLAVECAFRRRSVGI 387
DB 227 -----YRIGYE-TRITIGHVQGRGSPAFDRRLLSHGVEAVDALLDQEVDM 284

RESULT 9
T13433
pyrophosphate-dependent phosphofructo-1-kinase homolog T17A13.40 - Arabidopsis thaliana
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C.Accession: T13433
R.Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
submitted to the Protein Sequence Database, July 1999
A.Reference number: Z17683
A.Accession: T13433
A.Molecule type: DNA
A.Residues: 1-473 <BEV>
A.Cross-references: UNIPROT:Q9MOF9; EMBL:AL096692; GSPDB:GN00062; ATSP:T17A13.40
A.Experimental source: cultivar Columbia; BAC clone F17A13
C.Genetics:
A.Gene: ATSP:T17A13.40
A.Map position: 4

```


A.Introns: 35/3; 51/2; 64/3; 85/3; 131/3; 202/3; 225/3; 269/1; 358/3; 373/1; 424/3
 C/Superfamily: pyrophosphate-dependent phosphofructokinase, TP0108 type; 6-phosphofructo
 P.95-405/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 11.0%; Score 251; DB 2; Length 473;

Best Local Similarity 25.2%; Pred. No. 5e-10; Mismatches 174; Indels 72; Gaps 15;

Matches 103; Conservative 60; Mismatches 174; Indels 72; Gaps 15;

Qy 37 ILTAGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLGDSYPTAVERKKAGYL 96
 Db 97 ITCGGLCGMLNTVREIVCGLSYMGVGRILIGDSYGFVARNTHLDL---KTVNDI 153
 Qy 97 QPFGSVINSRVKLTNVKDCVKRGLVKEGDPQKVAADQLVKDVLIHTGGDDTNTA 156
 Db 154 HRSGGTILCTSR-----GGHNTTKI-VSIOQRGINQVYIIGGDSQXG 196
 Qy 157 AADLAFLARNYGLVIGLPTKVDNDVFPKISGLGAMTAEGAGYFNVAENANR 216
 Db 197 AALFEIRKRLKLVAVAGIPKTIIDIPIDISFGFPTDVAEQAIAAHVATSPEN 256
 Qy 217 MLIVHEVMGRNCGMLTAATAQERYKLLDBAEMLPBLGLTRESYEVAVFPEMAIDLEAE 276
 Db 257 GGLVLMKRYGFI-----WATLASRDVCCILPESPFLBGS 297
 Qy 277 A-----KRLBWDKVDVCNIFVSEGAVEAIVEMQAKGQEVPRDAFGHILDAVNP 329
 Db 298 GGLFEIRIDRLKESGHV-----IVIEGAGQDLISSEM---KESTTLKDSGNLQDI-- 349
 Qy 330 GKPFQGFQPMGAEKTLVOK---SGYFAPASNSNDVRLKSCADLAVECFRESSEV 386
 Db 350 GLMISRIKDHAKKWTLLTKTIDPTMTLRVAVNSND---NVCTLLAQSAVH---GV 402
 Qy 387 I-GHDEDNQNVLA---IEPPI-KGGKPFNIDDMFMSLSIGOP 428
 Db 403 MAGVNGFTGLVNGRHTYIPFNRIKOKKVVITDMMAR-LISSTNP 450

RESULT 10

J00016 6-phosphofructokinase (EC 2.7.1.11) alpha chain - yeast (Saccharomyces cerevisiae)

N.Alternate names: phosphofructokinase 1; phosphohexokinase; protein 194; protein G8599;
 C/Species: Saccharomyces cerevisiae
 C/Date: 31-Mar-1990 #sequence, revision 14-Jul-1994 #ext-change 09-Jul-2004
 C/Accession: J00016; S38963; S57708; S64566; S64564; S63924

R.Hellic, J.; Ritcel, R.G.; von Borestel, R.C.; Aguilera, A.; Rodicio, R.; Zimmermann,
 Gene 78, 309-321, 1989

A>Title: The phosphofructokinase genes of yeast evolved from two duplication events.

A/Reference number: A91608; UID:89378757; PMID:2528496

A/Accession: J00016

A/Molecule type: DNA

A/Residues: 1-987 <HEI>

A/Cross-references: UNIPROT:P16861; EMBL:M26943; NID:g172137; PIDN:AA34859.1; PID:g1721

R.Koperechleger, G.; Baer, J.; Stettinwagen, E.

Eur. J. Biochem. 217, 527-533, 1993

A>Title: Limited proteolysis of yeast phosphofructokinase. Sequence locations of cleavag

A/Reference number: S38963; UID:94039086; PMID:8223596

A/Accession: S38963

A/Molecule type: protein

A/Residues: 1-6; 90-97; 197-205; 914-921 <KOP>

R.van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.

submitted to the EMBL Data Library, June 1995

A/Description: Sequence analysis of the 43 KB CRM1-YLM9-PET54-SM11-PHO81-YHB4-PFK1 regio

A/Reference number: S57680

A/Accession: S57708

A/Molecule type: DNA

A/Residues: 794-987 <VAN>

A/Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA61193.1; PID:g886937

A/Experimental source: strain S288C

R.Guerrero, P.; Barreiro, T.; Azevedo, D.; Rodrigues-Pousada, C.

submitted to the Protein Sequence Database, May 1996

A/Reference number: S64565

A/Accession: S64566

A/Molecule type: DNA

A/Residues: 1-987 <GUB>

A/Cross-references: EMBL:E73025; NID:g1323434; PIDN:CAA97268.1; PID:g1323435; GSPDB:GN00C

A/Experimental source: strain S288C

R.van der Aart, Q.J.M.; Steensma, H.Y.

submitted to the Protein Sequence Database, May 1996

A/Reference number: S64541

A/Accession: S64564

A/Molecule type: DNA

A/Residues: 794-987 <VAN>

A/Cross-references: EMBL:E73025; GSPDB:GN00007; MIPS:YGR240C

A/Experimental source: strain S288C

R.van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.

Yeast 12, 385-390, 1996

A>Title: Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SM11-PHO81-YHB4-PFK1 region

A/Reference number: S63896; UID:96267763; PMID:8701610

A/Accession: S63924

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 794-987 <VAF>

A/Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA61193.1; PID:g886937

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C/Comment: Phosphofructokinase is composed of four alpha chains and four beta chains.

C/Genetics:

A/Genes: SGD:PFK1; MIPS:YGR240C

A/Cross-references: SGD:S0003472; MIPS:YGR240C

A/Map position: 7R

C/Superfamily: ATP-dependent phosphofructokinase, eukaryotic type; 6-phosphofructokinase

C/Keywords: ATP; glycolysis; phosphotransferase

F/208-517/Domain: 6-phosphofructokinase 1 homology <6PF1>

F/596-888/Domain: 6-phosphofructokinase 1 homology <6PF2>

F/325,329/Binding site: AMP, allosteric (Arg) #status predicted

F/356/Active site: Asp (Asp, Gly) #status predicted

F/398,455,488,491,729/Binding site: fructose-6-phosphate (Met, Glu, Lys, His, Arg, G)

F/751/Binding site: citrate, allosteric (Lys) #status predicted

Query Match 10.8%; Score 246.5; DB 1; Length 987;

Best Local Similarity 23.9%; Pred. No. 2.8e-09;

Matches 103; Conservative 68; Mismatches 135; Indels 125; Gaps 19;

Qy 31 KKKVAAILTAGLAPCLNSAIGSLIERYTEIDPSIEIICYRGYKGLLGDSYPTAVER 90
 Db 204 KKKKIAVMTSGDSDPKMNAVAVR--TGHPGCDVPAVYEGBELRGKRY----LK 256
 Qy 91 KKA-----GLRFGSGVIGNSRVKLTNVKDCVKRGLVKEGDPQKVAADQLVKDSVDIL 145
 Db 257 KKAEMEDVQWLS--GGTILGTA-----RSMERKREGRQAAGNLISGIDIL 303
 Qy 146 HTGGDDTNTAAADL-----AALFARNY-----GLTVIGLPTVNDVFPPI 187
 Db 304 VYCGGGSGLT--GADLFREHWPGLVDELVAEGFTBEVAPYKRLSIVGLVGSIDNDMSGT 362
 Qy 188 KQSLGAMTAEGGARYFMNVVAENANPMLIVHEVMGNCGLT---AATPOEYRKL 243
 Db 363 DSTIGVSLERICEVVDYIDATFASHSRAPV--EWMGHGNCMLMAGIATGADY---- 417
 Qy 244 DRAEMLPBLGLTRESYEVAVFVPEMAIDLBABAKRLREVMKVCV-----NIFPSGA 298
 Db 418 -----TIIPERAVPHGKODELSCORHRSKGRANNTIIVAEBA 457
 Qy 299 -----GVEAIVAEQKGOEVPDPAFGHKL---DAVNPQKMFGEQAPOMIAEK 345
 Db 458 LDDQLNPVTANVQKALIEL---GLDTKTYITIGHVQGGTAVAHBRWLAT--LQGVDAVK 512
 Qy 346 TLVQKSGYFARASASN-----VDMLRLKSCADLAVECAFRRSSGVIGHBE 391
 Db 513 AVLE-----FTPTSPSLGILENKTIIRMLVESVXLKTSVA-----TALENKPF 557
 Qy 392 DNGNVTARATEP 402
 Db 558 DKALISLDTTEP 568

RESULT 11
S32902
6-phosphofructokinase (EC 2.7.1.11) alpha chain - yeast (*Kluyveromyces marxianus* var. 1a
C)Species: *Kluyveromyces marxianus* var. *lactis*, *Candida sphaerica*
C)Date: 08-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C)Accession: S32902; S29602
R)Heinrich, J.; Kitchin, L.; Linsen, T.; Vogelberg, K.; Hollenberg, C.P.
Mol. Microbiol. 8, 559-570, 1993
A)Title: Molecular genetics of phosphofructokinase in the yeast *Kluyveromyces lactis*.
A)Reference number: S32902; MUID:93316853; PMID:8326866
A)Accession: S32902
A)Molecule type: DNA
A)Residues: 1-992 <HEI>
A)Cross-references: UNIPROT:003215, EMBL:Z17315, NID:g2879, P1DN:CAA78963.1; P1D:g2880
C)Genetics:
A)Gene: PFK1
C)Superfamily: ATP-dependent phosphofructokinase, eukaryotic type; 6-phosphofructokinase
C)Keywords: ATP, phosphotransferase
F:186-499/Domain: 6-phosphofructokinase 1 homology <6PF1>
F:1574-866/Domain: 6-phosphofructokinase 1 homology <6PF2>

Query Match 10.7%; Score 245.5; DB 2; Length 992;
Best Local Similarity 23.8%; Pred. No. 3.3e-09;
Matches 108; Conservative 74; Mismatches 165; Indels 107; Gaps 19;
QY 30 NKKPKVALITAGGLAPCLNSAIGSLIERYTEIDPSIEIICRYGKGLLDGSPVTAEV 89
DB 181 SKKKKIAVMTSGDGSQGMMAVRAVVR--SSIYGCDAVAAYEGEGLVKGDDYLRKMEW 238
QY 90 RKKAVALDRFGSGVIGNSRVKLTNVKDCVKGRLVEGED--PQKAAADQLVMDGVDILHT 147
DB 239 KQVRKMLSE--GSTLIGTA-----SKFRRMRKQKQCSNLDIGQIDALVV 283
QY 148 IGGDDTNTAA-----ADLAFLARN-----YGLTVIGLPTVDNVPIKOS 190
DB 254 IGGDSSLGADLFRSEWPSLVEIVKDKETFEVALYQNLITVGMWSIDMDSGTST 343
QY 191 LGAMTAABEGAGRYFMNVVAENNANPRLIVHEVMGNGCWL---TAATAQYRKLDPRA 246
DB 344 IGAVSALERICEMVDYIDATKASHRAFAV--EVMGRHCQWGLMSGIATADY----- 395
QY 247 EMLPRLGLTRESYEVHVAVPEMAI-----DLBAKRLREVDKVDVNIPIVSEGA- 298
DB 396 -----IFIPRAAPHGKMDQLKRVCCRHKEKRRNN--TVIAEGQL 436
QY 299 --GVEAIVAEQAK-----GOEVRDPAFGHIKL--DAVNPCKMFGEOFAQWIGAEKTLV 348
DB 437 DDQLNPITAE-QVKQVVELGIDTKITIGHVQROGTAVAHDRMLAT--LQGVDAVKAIL 493
QY 349 ---OKSGYFARASASVNDMLIKSCADLAVECAFRRSGVIGHEDENGVLRAIEPRR 404
DB 494 NMTPEPSPILGILNKVIRMPLEVES-----VKTRQVAAAIKADPFAKISLRDETFE 548
QY 405 I-----KGRPFNIDTWMFSMLSEIGOP 428
DB 549 LYSNFMSTTVDGSQLPEADRLNIAIVHGA 582

RESULT 12
T10691
probable diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90) T16118.50 -
N)Alternate names: protein T16118.50
C)Species: *Arabidopsis thaliana* (mouse-ear cress)
C)Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C)Accession: T10691
R)Beyan, M.; Lemnard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A)Reference number: Z16533
A)Accession: T10691
A)Molecule type: DNA
A)Residues: 1-462 <BEV>
A)Cross-references: UNIPROT:Q9M076, EMBL:AL049915, GSPDB:GM000662, ATSP:T16118.50
A)Experimental source: cultivar Columbia; BAC clone T16118

C)Genetics:
A)Gene: ATSP:T16118.50
A)Map position: 4
A)Introns: 35/3; 51/2; 64/3; 85/3; 131/3; 184/3; 202/3; 225/3; 269/1; 354/3; 371/1; 422/3;
C)Superfamily: pyrophosphate-dependent phosphofructokinase, T10108 type; 6-phosphofructo
C)Keywords: phosphotransferase
F:95-403/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 10.6%; Score 243; DB 2; Length 462;
Best Local Similarity 25.2%; Pred. No. 1.8e-09;
Matches 105; Conservative 58; Mismatches 178; Indels 76; Gaps 14;
QY 31 KPKKV--AIIAGGLAPCLNSAIGSLIERYTEIDPSIEIICRYGKGLLDGSPVTAEV 88
DB 89 KPSDRACTVCGGICPELNTVIREIVYGLHFMGVTEIVGDCGFRFPYSKNTALTP- 147
QY 89 VRKAGVLDKRGSGVIGNSRVKLTNVKDCVKGRLVEGEDPQKAAADQLVMDGVDILHT 148
DB 148 --KTVSDIHKGGITILGSR-----GHDTSKI--VDNIQDREINQVYII 188
QY 149 GGGDTNTAAADLAFLANNYGLTVIGLPTVDNVPIKOSLGAMTAABEGAGRYFMNV 208
DB 189 GGGDTQKGANIYKEIRRRGLKVAAGIPKTIIDNDIPVIDKSGFDVAEEQRAINAAH 248
QY 209 AENNANPRLIVHEVMGNGCWLTAATQAEYRKLDPRAEMLPGLTRESYEVHVAFVPE 268
DB 249 VEATSVENGIGIVKLMGRYSGFIA-----MYATLASDVIDCLLPE 289
QY 269 MAIDLEAE-----AKRLREVMKVDVNIPIVSEGAQVEAI---VAEMQAKGQEVPRDA 318
DB 290 SPFLIEGKGYEYFIARLRNGHNV-----IVIAEGQDILVBSIEQDASGNKLNDV 345
QY 319 FGHIKLDVNGKMGF-----EQPAQWIGAEKTLVQ--KSGYFARASASVNDMLIKSCAD 373
DB 346 -----GLMWSLKIEYFAKHNVMYDITLKIDPTVIMIRAIIPANASNVYSTLLAQ 394
QY 374 LAVECAPRRSGVIGHEDENGVLRAIEPRRI--KGRPFNIDTWMFSMLSEIGOP 428
DB 395 SAVGNMAGYGFVSGLVNGHTY--IPFNRIIRQNRKVITDRWMAR--MLSTNQ 448

RESULT 13
S54978
6-phosphofructokinase (EC 2.7.1.11), pyrophosphate-dependent - *Neisseria fowllei*
N)Alternate names: pyrophosphate-dependent phosphofructo-1-kinase
C)Species: *Neisseria fowllei*
C)Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
C)Accession: S54978; S71070
R)Wessberg, K.L.; Skolnick, S.; Xu, J.; Marciano-Cabral, F.; Kemp, R.G.
Biochem. J. 307, 143-149, 1995
A)Title: Cloning, sequencing and expression of the pyrophosphate-dependent phosphofructo-
A)Reference number: S54978; MUID:95234028; PMID:7717968
A)Accession: S54978
A)Molecule type: mRNA
A)Residues: 1-437 <MES>
A)Cross-references: UNIPROT:Q27705, EMBL:U11733
R)Kemp, R.G.
submitted to the EMBL Data Library, July 1994
A)Reference number: S71070
A)Accession: S71070
A)Molecule type: mRNA
A)Residues: 1-141, 'VTL', 145-437 <KEM>
A)Cross-references: EMBL:U11733, NID:9511230, P1DN:AA85791.1; P1D:9511231
C)Superfamily: pyrophosphate-dependent phosphofructokinase, T10289 type; 6-phosphofructo
C)Keywords: phosphotransferase
F:20-352/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 10.6%; Score 241.5; DB 2; Length 437;
Best Local Similarity 27.6%; Pred. No. 2.1e-09;
Matches 120; Conservative 65; Mismatches 171; Indels 79; Gaps 24;
QY 27 TLNKKPKV--AIIAGGLAPCLNSAIGSLIERYTEIDPSIEIICRYGKGLLDGSP 82

```

Db      9 TTITVPRKAVPTLGLVVGCGPAPGICNGVITGAATIE---AATNGRVULGFIEGPNLILDD 65
Qy      83 YPVTAEVRKKAAGVLOFGSGVIGNSRVKLTNNVCVKGGLVKEGEDPQKVAADQLVKDGV 142
      66 SKIVELTIDSVSRHIEGGSIILKTSRANPT-----KKQEDLQGV-VKQLCKFNV 113
Qy      143 DLHTTGGDDTTAAADLAFLARNYGLTVGLPQTVNDVFPKQSL---GAMTAAQ 199
      114 SLVYIGSD--TAFSGMSVAKAANN-ELHVCVPTINDL-PLPYGIPTEGYEARAEF 169
Db      200 GARYFNVAENNANPRMLIYHEVNGRNGMLT---ATAQERYKLDAEMLPGLR 256
      170 GANVVRNLTMTDASTASRYTV-VAMGRQAGHLALGKSAQSHLTIIPR-EFLP-----TT 223
Qy      257 ESVEVNAVE--VPEMAIDLEAE-AKRLREVMDKVCNIFVSEGAVEAIVAE--QAKQ 312
      224 DSTEPEVTFSRICDM---IEASIKRL--YTSKKDHGVILAEGLLEWSTDELKAFQS 278
Qy      313 EYPRDAFGHIKLDVAVNGKWFGEQFQMT---GAETLVQKS-GYPARASASNVDMRLI 368
      279 SLKYDAHDIMLAELDFGLVDEMEREMNRRLKIAFEKNLGYELRCAPVAFPREYT 338
Qy      369 KSCADAIVACAFRRSGVIGHDENGNVRAIEFPRIKGGK--PRNIDDMFNSMLSEIG 426
      339 RULGNAVRYLLANGGAL-----ITVQGVKNVPLSPD-----DLK 374
Db      427 QPKGK-----VEVS 436
      375 DPRTGKTRTROYDVS 389

```

RESULT 14

6-phosphofructokinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (ser

A/Species: Salmonella enterica subsp. enterica serovar Typhi

A/Note: this species has also been called Salmonella typhi

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004

C/Accession: AH0942

R.Parthill, J.; Dougan, G.; Jumea, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.J.; Moulé, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A/Author: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A/Reference number: AB0502; WUID:21534947; PMID:11677608

A/Accession: AH0942

A>Status: preliminary

A/Molecule type: DNA

A/Residues: 1-320 <PAR>

A/Cross-references: GB:AL513382; PIDN:CAD09562.1; PID:gl6504676; GSPDB:GN00176

C/Genetics:

A/Name: STY3809

C/Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase

Query Match 10.5%; Score 240.5; DB 2; Length 320;

Best Local Similarity 24.8%; Pred. No. 1.6e-09;

Matches 102; Conservative 59; Mismatches 122; Indels 129; Gaps 18;

33 KVAIITAGLAPCLNSAIGSLIERTEIDPSIEITCYRGYKGL-----LLGDSFPVTA 87

3 KRIGVLTSGDAPGMMAAIGVVR--AATTEGLEWGIYDYLGLYEDRMVQLDRYSVD 60

88 EVRKKAAGVLOFGSGVIGNSRVKLTNNVCVKGGLVKEGEDPQKVAADQLVKDGVILHT 147

61 MNR-----GGTILGSAAR--FPFPRD-----ENRAVAIEMLKKGKGIIDLVV 100

148 IGGDDTNTAAADLAFLARNYGLTVGLPQTVNDVFPKQSLGAMTAAEQGARYFNVA 207

101 IGGDGSYMAKRL-----TEMGFPCIGLPGITINDIKGDYIGFTALGTVAEIDRL 154

208 VAEENANPRMLIYHEVNGRNGMLTAAATQERYKLDAEMLPGLTRESYEVNAVP 265

155 RUTSSHORISIV-EVMGRYCGDLTAAA-----IAGGCEFIIVPEVERNR----- 200

```

Qy      266 VPEMAIDLEAEK-----RLREVMDKVCNIFVSEGAVE--AIVAEQAK 310
Db      201 -----DLVAEIKAGIAKKGKAIIVATEHMCDDVDLAFIEKEKTERETRAATLGH1QKG 254
Qy      311 GQVPRDAFGHIKLDVAVNGKWFGEQFQMTGAETLVQKSGYPARASASNVDMRLIS 370
      255 GSPVPD-----RLASRMGAYV----- 272
Db      371 CADLAVE-----CAFRSGVIGHDENGNVRAIEFPRIKGGKPPNIDTW 417
      273 -IDLLEIGHGRCVGIQNEQLVHND-----IIDALENMK---RPFK--SDW 312

```

RESULT 15

KIECPA

6-phosphofructokinase (EC 2.7.1.11) 1 - Escherichia coli (strain K-12)

N/Alternate names: phosphofructokinase 1, isozyme 1; phosphohexokinase, isozyme 1

C/Species: Escherichia coli

C/Date: 31-Dec-1988 #sequence_revision 10-Oct-1997 #text_change 09-Jul-2004

C/Accession: G65197; A25206; S40859

R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.

A/Reference number: A64720; WUID:97426617; PMID:9278503

A/Accession: G65197

A>Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-320 <BLAT>

A/Cross-references: UNIPROT:P06998; GB:AE000466; GB:U00096; NID:g2367328; PIDN:AACT6898.J

A/Experimental source: strain K-12, substrain MG1655

R.Helling, H.W.; Evans, P.R.

Eur. J. Biochem. 149, 363-373, 1985

A/Title: Nucleotide sequence and high-level expression of the major Escherichia coli pho

A/Reference number: A91144; WUID:85203917; PMID:3158524

A/Accession: A25206

A/Molecule type: DNA

A/Residues: 1-73; 'C', 75-102, 'DG', 105-162, 'P', 164-316, 'E', 318, 'M', 320 <HEL>

A/Cross-references: GB:X02519; NID:942365; PIDN:CAA26356.1; PID:g942366

A/Note: this sequence has since been corrected

R.Evans, P.R.

submitted to the EMBL Data Library, October 1986

A/Reference number: A94501

A/Content: annotation; corrections

R.Plunkett III, G.; Burland, V.; Daniels, D.L.; Blatner, F.R.

Nucleic Acids Res. 21, 3391-3398, 1993

A/Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region from 8;

A/Reference number: S40802; WUID:93347969; PMID:8346018

A/Accession: S40859

A>Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-320 <PLU>

A/Cross-references: EMBL:L19201; NID:g304961; PIDN:AAB03048.1; PID:g305019

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993

C/Comment: The active enzyme catalyzes the key control step of glycolysis, the phosphoryl

Y ADP and inhibited by phosphoenolpyruvate.

C/Comment: In E. coli this enzyme is responsible for 90% of the phosphofructokinase acti

ymes are different kinetically; there is also no immunological cross-reactivity.

C/Genetics:

A/Name: pfkA

A/Map position: 88 min

C/Superfamily: ATP-dependent phosphofructokinase, prokaryotic type; 6-phosphofructokinase

C/Keyword: allosteric regulation; ATP; glycolysis; homotetramer; phosphotransferase

F;3-219/Domains: 6-phosphofructokinase 1 homology <6PPI>

Query Match 10.4%; Score 238.5; DB 1; Length 320;

Best Local Similarity 24.9%; Pred. No. 2.2e-09;

Matches 100; Conservative 65; Mismatches 132; Indels 105; Gaps 18;

33 KVAIITAGLAPCLNSAIGSLIERTEIDPSIEITCYRGYKGL-----LLGDSFPVTA 87

3 KRIGVLTSGDAPGMMAAIGVVR--AATTEGLEWGIYDYLGLYEDRMVQLDRYSVD 60

```

QY      88 EVRKXAGVUQRFGSVIGNSRYKLTIVKDCVKGGLVKEGEDPQKVAADQLVKQGVDIJHT 147
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 MINR-----GGTFLGSR--FPEFRD-----ENIRAVAIENLKRRGIDALVV 100
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      148 IGGDDTNTAAADLAAFLAARNNYGLTVIGPKTVNDVFPPIKOSLGAWTAEOGARVFMV 207
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      101 IGGDGSYMGAMRL-----TEMGFPCIGLPGTIIDNDIKGTDTYTIIGFTALSTVVEAIDRL 154
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      208 VAENNANPRLIVHEVWGNCGMLTAATAOEYRKLDRAEW--LPELGLTRESYEVHAVF 265
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      155 RDTSSSHQRISYV-EVMGRYCGDLTLAA-----IAGGCEFYVVEVEFSRED-----L 202
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      266 VPEMAIDLEAEAKR-----LREVMKVDCVNI FVSEGAVE--AIVAEQAKQOEVPFRD 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      203 VNEIRAGI-AKGRKAIVAITEHMCVDDELAFIEKETRETRATVTLGHIOGGSPVPYD 261
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      318 AFGHITKLDAVNPCKMFGEOFAQMIGAEKTLVQSGYFARASASNVDMRLIKSCADLAVE 377
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      262 RI-----LASRMGAYAIIDLLAGYGR----- 283
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      378 CAFRRBSGVIGHDEDNQNVLRALIEPPRIKGGKPFNIDTDFN 419
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      284 CVGIQNEQLVHHD-----IIDAIEENMK---RPFK--GDWLD 314
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: February 24, 2005, 14:23:20
 Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 08:27:31 ; Search time 5943 Seconds
(without alignments)
3563.003 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTWPHYHATDIRCHNMF.....FNSMLSEIQPKXGKGVESH 437

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=xlh
-O=/gen2.1/USPTO_spool/US10701200/runat_25022005_102832_10310/app_query.fasta_1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=DCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=DCT -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10701200 @GCM 1 1 3731 @runat_25022005_102832_10310 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEUDERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONCLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_gy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|------|---------------------|
| 1 | 2284 | 100.0 | 1311 | 6 AR308857 Sequence |
| 2 | 2284 | 100.0 | 1311 | 6 AR473868 Sequence |
| 3 | 2284 | 100.0 | 1311 | 6 AR565849 Sequence |
| 4 | 2284 | 100.0 | 1311 | 6 AX394958 Sequence |

| | | | | |
|----|--------|-------|--------|----------------------|
| 5 | 2284 | 100.0 | 1311 | 6 AX398319 Sequence |
| 6 | 2284 | 100.0 | 1311 | 6 AX398428 Sequence |
| 7 | 1531.5 | 67.1 | 295650 | 1 BX294151 Pirellula |
| 8 | 1453 | 63.6 | 1427 | 3 AF246209 Mastigamo |
| 9 | 1444.5 | 63.2 | 8530 | 1 CQ363740 Sequence |
| 10 | 1434.5 | 62.8 | 110000 | 1 AB017283_11 |
| 11 | 1314 | 57.5 | 1400 | 1 PRSPFR |
| 12 | 1302 | 57.0 | 294800 | 1 SME591789 |
| 13 | 1292.5 | 56.6 | 11715 | 1 AE008127 |
| 14 | 1292.5 | 56.6 | 11807 | 1 AE009161 |
| 15 | 311 | 13.6 | 1549 | 1 SCU51728 |
| 16 | 311 | 13.6 | 321250 | 1 SC0939111 |
| 17 | 305 | 13.4 | 1233 | 1 AF268276 |
| 18 | 298.5 | 13.1 | 2302 | 1 AMU31277 |
| 19 | 298 | 13.0 | 299925 | 1 AP005045 |
| 20 | 284.5 | 12.5 | 299425 | 1 AP005049 |
| 21 | 281.5 | 12.3 | 1458 | 8 AY040055 |
| 22 | 281.5 | 12.3 | 2087 | 8 AF360207 |
| 23 | 281.5 | 12.3 | 298050 | 1 AP003189 |
| 24 | 281.5 | 12.3 | 339650 | 1 SC0939108 |
| 25 | 279.5 | 12.2 | 110000 | 1 AE016822_07 |
| 26 | 279 | 12.2 | 141027 | 8 AC132486 |
| 27 | 279 | 12.2 | 163570 | 8 AC130730 |
| 28 | 275 | 12.0 | 110000 | 1 CP000013_7 |
| 29 | 274.5 | 12.0 | 2490 | 1 AF36847 |
| 30 | 272.5 | 11.9 | 110000 | 2 AP006495_4 |
| 31 | 271.5 | 11.9 | 110000 | 1 AE017180_22 |
| 32 | 271.5 | 11.9 | 300100 | 1 SC0939123 |
| 33 | 270 | 11.8 | 1778 | 6 AR050093 |
| 34 | 270 | 11.8 | 11813 | 1 AE001172 |
| 35 | 269.5 | 11.8 | 281450 | 1 AP005032 |
| 36 | 267.5 | 11.7 | 300205 | 1 AE015944 |
| 37 | 266 | 11.6 | 1467 | 8 BT015466 |
| 38 | 266 | 11.6 | 1470 | 8 AY090267 |
| 39 | 266 | 11.6 | 1990 | 8 AY049245 |
| 40 | 264.5 | 11.6 | 1624 | 6 AR050094 |
| 41 | 263.5 | 11.5 | 1327 | 3 EHPPIPEK |
| 42 | 263.5 | 11.5 | 2256 | 6 AX654108 |
| 43 | 262 | 11.5 | 1978 | 6 AR050092 |
| 44 | 262 | 11.5 | 15158 | 1 AE001705 |
| 45 | 261.5 | 11.4 | 1490 | 6 AX654128 |

ALIGNMENTS

RESULT 1
AR308857
LOCUS AR308857 1311 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 15 from patent US 6555353.
ACCESSION AR308857
VERSION AR308857.1 GI:31700588
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1311)
AUTHORS Koffas, M., Norton, K.C., Odom, J.M. and Ye, R.W.
TITLE Methanotrophic carbon metabolism pathway genes and enzymes
JOURNAL Patent: US 6555353-A 15 29-APR-2003;
FEATURES
source 1..1311
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,12e-161
Score: 2284.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6
Length: 1311
Matches: 437
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-701-200-6 (1-437) x AR308857 (1-1311)

QY 1 AapValThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe 20
 Db 1 GATGGTCACATGAGCCCTATCATTAACGGCTGATATCGATTGTGATCGATTTT 60
 QY 21 LeuAnpPheAnpPheTyrThrLeuMetAsnLysProLysLysValAlaIleLeuThrAla 40
 Db 61 CTTAACTTTAACTTCTACACGCTCATGAACCAACCTAAAGATTGCAATCTGACAGA 120
 QY 41 GlyGlyLeuAlaProCysLeuAnpSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
 Db 121 GCGCGCTTGCGCCCTTGTGTAATCCGAATCCGTAAGTTGATCGAAGCTTATACCA 180
 QY 61 IleAapProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
 Db 181 ATCGATCTCAGCATGAAATCATTTGCTATCGCGGGGTTATAAGGCGCTGTGCGG 240
 QY 81 AapSerTyrProValThrAlaGluValArgLysLysValGlyValLeuGlnArgPheGly 100
 Db 241 GATTTCTTATCCAGTAACGGCCGAAAGCGTAAAGCGGGGTGTTCTGCAAGCTTTGGC 300
 QY 101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAapCysValLysArg 120
 Db 301 GGTTCGTGATCGGCAACAGCCCGCTCAATGACCAATGCTCAAGACTGCTGAAAGC 360
 QY 121 GlyLeuValLysGlyGlyGluAapProGlnLysValAlaIleAapGlnLeuValLysAap 140
 Db 361 GGTTCGTCAAGAGGGTGAAGATCGCAAAAGTCGCGCGCTGATCAATTTGGTTAAGAT 420
 QY 141 GlyValAapIleLeuHisThrIleGlyLysAapThrAsnThrAlaAlaIleAapLeu 160
 Db 421 GGTTCGATATCTGCACACCATCGCGCGGATGATACCAATACGACGACGCGGATTTG 480
 QY 161 AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrVal 180
 Db 481 CGAGCATCTCTGGCCGAAATTAATTCGACGACGCTGATGTTTACTTAAACCGTC 540
 QY 181 AapAnpAapValPheProIleLysGlnSerLeuGlyValaTrpThrAlaAlaGlnGlnGly 200
 Db 541 GATTAACGACGATATTCGATCAAGCAATCATAGGTGCTTGACGCGCCGACGAAAGC 600
 QY 201 AlaArgTyrPheMetAsnValValaIleGluAsnAnpAlaAnpProArgMetLeuIleVal 220
 Db 601 GCGCGTATTTATGATGAACGTGGTGGCGCAAAACAACGCAACCCACGCACTGATCGTA 660
 QY 221 HisGluValMetGlyArgAsnCysGlyTyrLeuThrAlaAlaIleThrAlaGlnGlnTyrArg 240
 Db 661 CACGAAGTATGAGCGCGCTAATCGCGCTGCTGACCGCTGCAACCGCGAGGAAATTCGC 720
 QY 241 LysLeuLeuAapArgAlaGluTrpLeuProGluLeuGlyLeuThrArgGlnSerTyrGlu 260
 Db 721 AAATTAACCTGACCGTGGCCGAGTGGTGGCGGAATGGGTTGACTCGTAATCTTATGAA 780
 QY 261 ValHisAlaValPheValProGluMetAlaIleAapLeuGluAlaGluAlaLysArgLeu 280
 Db 781 GTCACGCGGATTTGCTTCCGAAATGGGATCGACTGGAAGCCGAAGCCGAAGCGCTCG 840
 QY 281 ArgGluValMetAspLysValaIleAapCysValaIlePheValSerGlnGlyAlaGlyVal 300
 Db 841 CCGGAAGTATGACCAAGTCAATTCGCTCAACATCTTCTCCGAAGTCCCGCGCTC 900
 QY 301 GluAlaIleValaIleGluMetGlnAlaLysGlyGlnGluValProArgAapAlaPheGly 320
 Db 901 GAAGCTATGTCGCGAAATGACAGGCGCAAGGCGCAAGTCCGCGCATCGTTCCGCGC 960
 QY 321 HisIleLysLeuAapAlaValaIleAnpProGlyLysTrpPheGlyGlnGlnPheAlaGlnMet 340
 Db 961 CACATCAAACTGAATGCGGTCAACCTCGTAAATGGTTTCGGGACCAATTCGCGCAGATG 1020
 QY 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 Db 1021 ATAGCGCGGAAAAACCTGATCAAAAATCGGAGTACTTTCGCGCGTCTTCTGCTTCC 1080

QY 361 AsnValAapAapMetArgLeuIleLysSerCysAlaAapLeuAlaValGluCysAlaPhe 380
 Db 1081 AACGTTGACGACATGCGTTTATCAAAATCGTCCGCGCTTGGCGGTGAGGCGCGCTTC 1140
 QY 381 ArgArgGlnSerGlyValIleGlyHisAapGluAapAnpGlyAsnValLeuArgAlaIle 400
 Db 1141 CCGCGAGTCTGGGGTATCGTTCACACGAAAGCAACGCGCAACGCTTGGCTCGATC 1200
 QY 401 GluPheProArgIleLysGlyGlyLysProPheAnpIleAapThrAapTrpPheAnpSer 420
 Db 1201 GAGTTCCGCCCATCAAGGGGGGGAACCGTTCAATATCGACACCGACTGTTCAATAGC 1260
 QY 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
 Db 1261 ATGTTGACGGAATCGGCGCACGCTAAAGCGGTAAAGTCGAAGTCAGCGCAC 1311
 RESULT 2
 AR473868
 LOCUS AR473868 1311 bp DNA linear PAT 20-FEB-2004
 DEFINITION Sequence 5 from patent US 6689601.
 ACCESSION AR473868
 VERSION AR473868.1 GI:42712367
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1311)
 AUTHORS Koffas, M., Odom, J.M., and Schenzle, A.
 TITLE High growth methanotropic bacterial strain
 JOURNAL Patent: US 6689601-A 5-10-FEB-2004;
 FEATURES
 source
 1. .1311
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,12e-161 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Db: 6 Gaps: 0
 US-10-701-200-6 (1-437) x AR473868 (1-1311)
 QY 1 AapValThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe 20
 Db 1 GATGGTCACATGAGCCCTATCATTAACGGCTGATATCGATTGTGATCGATTTT 60
 QY 21 LeuAnpPheAnpPheTyrThrLeuMetAsnLysProLysLysValAlaIleLeuThrAla 40
 Db 61 CTTAACTTTAACTTCTACACGCTCATGAACCAACCTAAAGATTGCAATCTGACAGA 120
 QY 41 GlyGlyLeuAlaProCysLeuAnpSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
 Db 121 GCGCGCTTGCGCCCTTGTGTAATCCGAATCCGTAAGTTGATCGAAGCTTATACCA 180
 QY 61 IleAapProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
 Db 181 ATCGATCTCAGCATGAAATCATTTGCTATCGCGGGGTTATAAGGCGCTGTGCGG 240
 QY 81 AapSerTyrProValThrAlaGluValArgLysLysValGlyValLeuGlnArgPheGly 100
 Db 241 GATTTCTTATCCAGTAACGGCCGAAAGCGTAAAGCGGGGTGTTCTGCAAGCTTTGGC 300
 QY 101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAapCysValLysArg 120
 Db 301 GGTTCGTGATCGGCAACAGCCCGCTCAATGACCAATGCTCAAGACTGCTGAAAGC 360
 QY 121 GlyLeuValLysGlyGlyGluAapProGlnLysValAlaIleAapGlnLeuValLysAap 140

361 GGTTCGTCAGAAAGGCGTGAAGATCCGCAAAAAGTCGGCGCTGATCAATGGTTAAGAT 420
141 G1yValAsp11eleuH1sthr11leg1y1yAspAspThrAsnThrAla1AlaAspLeu 160
421 GGTCGTGATATTCGACACACATCCGCGCGGATGATACCAATACGCGACAGCGGATTTG 480
161 AlAlaPheLeuAlaArgAsnAsnTyrG1yLeuThrVal11leg1yLeuProlyThrVal 180
481 GCAGCATTCCTGGCCGAGAAATATTAATACGATGACCGATGTTTACTTAAACCGTC 540
181 AspAsnAspValPheProT1elysG1nserLeuG1yAla1TrpThrAla1AlaG1nG1y 200
541 GATTAACGATATTCCTGATCAAGCAATCACTAGGTGCTGATCCCGCGAGCAAGGC 600
201 AlaArgTyrPheMetAsnVal1AlaG1uAsnAsnAlaAsnProArgMetLeu11Ala 220
601 GCGCGTATTTTCAATGAACGTGGTGGCGAAACAGCCCAACCCAGCATGCTGATGTA 660
221 H1sG1uVal1MetG1yArgAsnCy8G1yT1rPLeuThrAla1AlaThr1AlaG1nG1y 240
661 CACGAAGTATGGCGCGTAACTGCGGCTGCTGACCGCTGCAACCGCGAGAAATACGC 720
241 LysLeuLeuAspArg1AlaG1uT1rPLeuProG1uLeuG1yLeuThrArg1uSerTyrG1u 260
721 AAATTACTGACCGTCCAGATGCTGCGGAATGGGATTGACTGCTGATCTTATGA 780
261 ValH1sAlaValPheVal1ProG1uMetAla1LeAspLeuG1uAlaG1uAlaValArgLeu 280
781 GTGACGCGGTATTCCTTCGGAATGGGATGATGATCGATCGAACCCAGACCAAGCCCTG 840
281 ArgG1uVal1MetAspLysValAspCy8Val1Asn11PheValSerG1nG1yAlaG1yVal 300
841 CCGGAAGTATGAGCAAAAGTCATGCGATCGATCAATCTTCGTTCCGAGGTGCGCGTC 900
301 G1uAla11LeuAla1AlaG1uMetG1nAla1yG1yG1uG1uVal1ProArgAspAlaPheG1y 320
901 GAAGCTATGCTCGCGAAATGACAGGCCAAAGCCAAAGTGGCGCGATGCTGCTGCGC 960
321 H1s11LeuLeuAspAlaVal1AsnProG1yLysT1rPheG1yG1uG1nPheAlaG1Met 340
961 CACATTAATCTGATCGGTCACCTGTTAATGTTCCGCGACCAATTCGCGCAGATG 1020
341 11leg1yAlaG1uLysThrLeuVal1G1uLysSerG1yTyrPheAlaArg1AsnAlaSer 360
1021 ATAGGCGCGGAAAAAACCTCGTGAACAAATCGGGAATCTTCGCCGCTGCTTCGCTTC 1080
361 AsnVal1AspAspMetArgLeu11LysSerCy8AlaAspLeuAlaVal1G1uCy8AlaPhe 380
1081 AACGTTGACGACATGCGTTTGATCAATCGTGGCGCATTTGGCGGTGCGATGCGGCTTC 1140
381 ArgArgG1uSerG1yVal11leg1yH1sAspG1uAspAsnG1yAsnVal1LeuArgAla11e 400
1141 CCGCGGAGCTGCGCGTGAATCGGTCACGAGAAAGACACGCAACGTTGCGTGGATC 1200
401 G1uPheProArg11elysG1yG1yLysProPheAsn11eAspThrAspT1rPheAsnSer 420
1201 GAGTTTCCCGCATCAAGGCGGCAACGCTTCAATATGACACCGACTGTTCAATAGC 1260
421 MetLeuSerG1u11leg1yG1nProlyG1yG1yVal1G1uVal1SerHis 437
1261 ATGTTGAGCGAATCGCGCAGCTTAAAGGCGGTAAAGTCGAAGTCAGCCAC 1311

RESULT 3
ARS65849 1311 bp DNA linear PAT 08-OCT-2004
LOCUS ARS65849
DEFINITION Sequence 15 from patent US 6767744.
ACCESSION ARS65849
VERSION ARS65849.1 GI:53981913
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1311)

AUTHORS Koffas, M., Norton, K.C., Odum, J.M. and Ye, R.W.
TITLE Methanotrophic carbon metabolism pathway genes and enzymes
JOURNAL Patent: US 6767744-A 15 27-JUL-2004;
FEATURES Location/Qualifiers
source 1..1311
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2,126-161 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-701-200-6 (1-437) x ARS65849 (1-1311)

QY 1 AspVal1ThrTrpProTyrH1sLeuThrAlaAsp11eAspPheCy8H1sTrpPhe 20
DB 1 GATGTGTCACATGGCCCTATCACTTAAGCGGTGATTCGATTTGTCACTTGGTTTTT 60
QY 21 LeuAsnPheAsnPheTyrThrLeuMetAsnLysProLysVal1Ala11eleuThrAla 40
DB 61 CTTAACTTAACTTCAACAGCTCATGAACAAACCTTAAAGTTGCAATCACTGACAGCA 120
QY 41 G1yG1yLeuAlaProCy8LeuAsnSerAla11leg1ySerLeu11leg1yArgTyrThrG1u 60
DB 121 GCGGCTTGCGCTTGTGTAATTCGCAATCGGTGATGATGCAACCTTATACCGAA 180
QY 61 11eAspProSer11leg1y11e11eCy8TyrArgG1yG1yTyrLysG1yLeuLeuG1y 80
DB 181 ATGATCTTACATTAATAATCATTTCTATCCGCGCGTTATTAAGCCCTTGCTGCGC 240
QY 81 AspSerTyrProVal1Thr1AlaG1uVal1ArgLysLysAlaG1yVal1LeuG1nArgPheG1y 100
DB 241 GATCTTATTCAGTAAACGCCCGAAGTCGTAAAGCGCGGTGTTCTGCAACGTTTGGC 300
QY 101 G1ySerVal111leg1yAsnSerArgVal1LysLeuThrAsnVal1LysAspCy8Val1Arg 120
DB 301 GGTTCGTATCGGCAACAGCGCGTCAATTAATGACCAATGTCAAAGCTGCGGAAACGC 360
QY 121 G1yLeuVal11leg1yG1uG1uAspProG1nLysVal1Ala1AlaAspG1nLeuVal1LysAsp 140
DB 361 GGTTCGTCAAGAGGCTGAAGATCCGCAAAAGTGGCGCTGATCAATTCGTTAAGAT 420
QY 141 G1yVal1Asp11eleuH1sthr11leg1y1yAspAspThrAsnThrAla1AlaAspLeu 160
DB 421 GGTCGATATTCGACACACATCGCGCGGATGATACCAATACCGCAGCGGATTTG 480
QY 161 AlAlaPheLeuAlaArgAsnAsnTyrG1yLeuThrVal11leg1yLeuProlyThrVal 180
DB 481 GCAGCATTCCTGGCCGAGAAATATTAATACGATGACCGTCAATGTTTACTTAAACCGTC 540
QY 181 AspAsnAspValPheProT1elysG1nserLeuG1yAla1TrpThrAla1AlaG1nG1y 200
DB 541 GATTAACGATATTCCTGATCAAGCAATCACTAGGTGCTGATCCCGCGAGCAAGGC 600
QY 201 AlaArgTyrPheMetAsnVal1AlaG1uAsnAsnAlaAsnProArgMetLeu11Ala 220
DB 601 GCGCGTATTTTCAATGAACGTGGTGGCGAAACAGCCCAACCCAGCATGCTGATGTA 660
QY 221 H1sG1uVal1MetG1yArgAsnCy8G1yT1rPLeuThrAla1AlaThr1AlaG1nG1y 240
DB 661 CACGAAGTATGGCGCGTAACTGCGGCTGCTGACCGCTGCAACCGCGAGAAATACGC 720
QY 241 LysLeuLeuAspArg1AlaG1uT1rPLeuProG1uLeuG1yLeuThrArg1uSerTyrG1u 260
DB 721 AAATTACTGACCGTCCAGATGCTGCGGAATGGGATTGACTGCTGATCTTATGA 780
QY 261 ValH1sAlaValPheVal1ProG1uMetAla1LeAspLeuG1uAlaG1uAlaValArgLeu 280

Db 781 GTGACGCGGTTTCTTCGCGAAATGGCGATCGACTTGGAAGCCGAAAGCCGCTG 840
 Qy 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyValIleVal 300
 Db 841 CGCGAAGTGAATGACAAAGTCGATTGCGGTCACATCTTCGTTCCGAAAGTCCGCGCTC 900
 Qy 301 GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 Db 901 GAAGCTATCGTCGCGAAATGCAAGCCAAAGCCGAGAAAGTCCGCGCATCTGTTCCGC 960
 Qy 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 Db 961 CACATCAAACTGATGCGCGTCAACCTCGTAAATGCTTCGCGACCAATTCGCGCATG 1020
 Qy 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 Db 1021 ATAGCGCGGAAAAAACCTGGTACAAAATCGGGATCTTCGCGCGTCTTCGCTTCC 1080
 Qy 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
 Db 1081 AACGTTGACGACATCGTTGATCAAAATCGTCGCGCATTCGCGCATCGATCGCGCTTC 1140
 Qy 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
 Db 1141 CGCCGCGAGTCTGCGCTGATCGGTCAACGAAAGCAACGCGCAAGCTGTTCGCTCGATC 1200
 Qy 401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspThrAspTrpPheAsnSer 420
 Db 1201 GAGTTTCGCGCGCATCAAGGCGGCAACCGTTCAATATGACACCGAAGTGTTCATATGC 1260
 Qy 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
 Db 1261 ATGTTGACGAAATCGCGCAAGCCCTAAAGCGGTAAAGTCGAAGTCAGCCAC 1311
 RESULT 4
 AX394958
 LOCUS AX394958 1311 bp DNA linear PAT 18-MAY-2002
 DEFINITION Sequence 1 from Patent WO0218617.
 ACCESSION AX394958
 VERSION AX394958.1 GI:21066031
 KEYWORDS
 SOURCE
 ORGANISM
 Methylobionas sp.
 Methylobionas sp.
 Bacteria; Proteobacteria; Gammaproteobacteria; Methylobionas;
 Methylobionasaceae; Methylobionas.
 REFERENCE
 1 Brzostowicz, P.C., Cheng, Q., Dicosimo, D.J., Koffas, M., Miller, E.S.,
 Odom, J.M., Plicataggio, S.K. and Rouviere, P.E.
 Carotenoid production from a single carbon substrate
 Patent: WO 0218617-A 1 07-MAR-2002;
 E.I. DUPONT DE NEMOURS AND COMPANY (US)
 FEATURES
 source 1..1311
 /organism="Methylobionas sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:418"
 /note="16a"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,126-161 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-701-200-6 (1-437) x AX394958 (1-1311)
 Qy 1 AspValAlaThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhePhe 20
 Db 1 GATGTGTCACATGCGCCCTATCACTTAACGGCTGATATTCGATTGTCATGCTTTT 60

Qy 21 LeuAsnPheAsnPheTyrThrLeuMetAsnLysProLysValAlaIleLeuThrAla 40
 Db 61 CTTAACTTAACTTCTTACACCTCATGAACAACTTAAAGTTGCAATATCGACAGA 120
 Qy 41 GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
 Db 121 GCGGCTTGCGGCTTGTGTTGTAATTCGCGAATCGGTAGTTGATGAAAGCTTATCCGA 180
 Qy 61 IleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
 Db 181 ATCATCTTACATATGAATATCATTTGCTATCGCGCGGTTATTAAGCCCTGTGCTGGGC 240
 Qy 81 AspSerTyrProValThrAlaGluValArgLysValAlaGlyValLeuGlnArgPheGly 100
 Db 241 GATCTTATCCAGTAAACCGCGAAGTGGTAAAGGGCGGTCTCGCAAGCTTTGGC 300
 Qy 101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
 Db 301 GGTTCGTGATCGGCAACGCGCGCTCAAAATTGACCAATGTCAAAGATCGCGTAAAGCC 360
 Qy 361 GlyLeuValLysGluGlyLysAspProGlnLysValAlaAspGlnLeuValLysAsp 140
 Db 421 GGTTCGTCAATATTCGACACCATCGCGCGCATGATACCAATACGCGACGCGAATTTG 480
 Qy 481 AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrVal 180
 Db 541 GATACGACGATATTCGATCAAGCAATCATAGGTGCTGAGTACGCGCGACGAGAGCC 600
 Qy 601 AlaArgTyrPheMetAsnValAlaAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
 Db 221 HisGluValMetGlyArgAsnCysGlyTyrLeuThrAlaAlaThrAlaGlnGluTyrArg 240
 Qy 241 LysLeuLeuAspArgAlaGluTyrLeuProGluLeuGlyLeuThrArgGluSerTyrGlu 260
 Db 721 AAATTACTGACCGCGGAGGTGTCGCGAATGGGTTTACATCGTCAATCTTATGAA 780
 Qy 781 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaLysArgLeu 280
 Db 841 GTGACGCGGATTCGTTCCGAAATGCGATCGACTTGGAAGCCGAAGCCAAAGCGCCTG 840
 Qy 841 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyValIleVal 300
 Db 901 CGCGAAGTGAATGACAAAGTCGATTGCGTCAACATCTTCGTTCCGAAGTCCGCGCTC 960
 Qy 901 GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 Db 961 GAAGCTATCGTCGCGAAATGCAAGCCAAAGCCGAGAAAGTCCGCGCATCTGTTCCGC 960
 Qy 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 Db 1021 ATAGCGCGGAAAAAACCTGGTACAAAATCGGGATCTTCGCGCGTCTTCGCTTCC 1080
 Qy 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 Db 1081 AACGTTGACGACATCGTTGATCAAAATCGTCGCGCATTCGCGCATCGATCGCGCTTC 1140
 Qy 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400

Db 1141 CCCCCGAGCTCGCTGATCGGTCAAGCAAGACAGCGCAACCTGTTGCGTGGATC 1200
 Oy 401 GIupheProxArgIlelysgIylybProphesAnIleasPhrThraPheAnSer 420
 Db 1201 GAGTTTCGCGCATCAAGGACGGCAACCGTTCAATATGACACCACTGGTTCAATAGC 1260
 Oy 421 MetLeuSerGluIleGIyGlnProlysgIylybValGluValSerHis 437
 Db 1261 ATGTTGAGGAATCGGCAAGCCTTAAGCGGTAAGTGAAGTCAAGTCAAGCAC 1311
 RESULT 5
 AX398319 1311 bp DNA linear PAT 27-MAY-2002
 LOCUS Sequence 15 from Patent WO0220796.
 DEFINITION AX398319
 ACCESSION AX398319
 VERSION AX398319.1 GI:21261097
 KEYWORDS
 SOURCE
 ORGANISM
 Methylobionas sp.
 Bacteria; Proteobacteria; Gammaproteobacteria; Methylcoccales;
 Methylcocccaceae; Methylobionas.
 REFERENCE
 1 Koffas,M., Odum,J.M., Norton,K.C. and Ye,R.W.
 Methanotrophic carbon metabolism pathway genes and enzymes
 Patent: WO 0220796-A 15 14-MAR-2002;
 E.I. DUPONT DE NEMOURS AND COMPANY (US)
 FEATURES
 source
 1..1311
 Location/Qualifiers
 -/organism="Methylobionas sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:418"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,126-161 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-701-200-6 (1-437) x AX398319 (1-1311)
 Oy 1 AsnValValThrThrProGlyHisLeuThrAlaAspIleAspPheCysHisTrpPhePhe 20
 Db 1 GATGGTGCACATGGCCCTATCACTTAACGGCTGATATTCGATTTGTCATTTGGTTTTT 60
 Oy 21 LeuAsnPheAsnPheTyThrLeuMetAsnLysProlybValAlaIleLeuThrAla 40
 Db 61 CTTAACTTTAACTTCAACGCTCATGAACCAACCTTAAGAAAGTTGCATTAAGACAGCA 120
 Oy 41 GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyThrGlu 60
 Db 121 GCGGGCTGGCGCTTGTGAATTCGCAATCGGATGTTGATCAACGTTATACGGA 180
 Oy 61 IleAsnProSerIleGluIleIleCysTyThrArgGlyGlyTyThrGlyLeuLeuLeuGly 80
 Db 181 ATCGATCTTAAGCAATTAAGATTCATTCGCGCGGTTATTAAGCCCTGTGTCTGGGG 240
 Oy 81 AspSerTyProValThrAlaGluValArgGlyLysAlaGlyValLeuGlnArgPheGly 100
 Db 241 GATTTCTTAATCCAGTAACGGCCGAAGTGGCTTAAGAAAGCGGGTCTTTCGCAAGTTTGGG 300
 Oy 101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
 Db 301 GGTTCGTGATCGGCAACAGCCCGTCAATTCACCAATGTCAAAGACTGCGTGAACCG 360
 Oy 121 GlyLeuValLysGluGlyAspProGlnLysValAlaAlaAspGlnLeuValLysAsp 140
 Db 361 GGTTCGTGCAAAAGGATGAGATCCGCAAAAGTGGCGGCTGATCAATTTGGTTAAGAT 420
 Oy 141 GlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeu 160

Db 421 GGTGTGATATTTCTGCACACCAATCCGCGCGATGATACCAATACGACAGCGGATTTG 480
 Oy 161 AlaAlaPheLeuAlaArgAsnAsnTyGlyLeuThrValIleGlyLeuProlybSerVal 180
 Db 481 GAGCATTTCTGGCCAGAAATTAATTAAGCATGACCGTCAATGGTTTACTTAAGACGTC 540
 Oy 181 AsnAsnAspValPheProIleLybGlnSerLeuGlyAlaTrpThrAlaIleGluGlnGly 200
 Db 541 GATTAACGATATTTCTCCATCAAGCAATCATAGTGGTTCGATCCCGCGAGCAAGGC 600
 Oy 201 AlaArgTyPheMetAsnValAlaAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
 Db 601 GCGCGTTATTTATCATGAACGTGGTGGCCGAACCAACGCCAACCCAGCATGCTGATCGTA 660
 Oy 221 HisGluValMetGlyValArgAsnCybGlyTyTrpLeuThrAlaIleThrAlaGlnGlyTyArg 240
 Db 661 CACGAAGTATGGCGCGTAACCTGGCTGCTGACCGCTGCAACCGCGAGGAATATCGC 720
 Oy 241 LysLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArgGluSerTyArg 260
 Db 721 AATTTACTGAGCCGTCCGAGTGTTCGCGGAATGGTTGACTGTGATCTTATGAA 780
 Oy 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu 280
 Db 781 GTGCACGCGGTATTCGTTCCGGAATGGCGATCGACCTGGAAGCCGAAGCCAGCCCTG 840
 Oy 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGlyAlaGlyVal 300
 Db 841 CCGGAAGTATGAGCAAAATGCAATTCGTCACATCTTGTTCCGAAGTGGCGCGTC 900
 Oy 301 GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 Db 901 GAAGCTATCGTCGCGGAATATGACAGGCCAAGCGCAAGAGTGGCGCGCATGCTTCGCG 960
 Oy 321 HisIleLysLeuAspAlaValAsnProGlyTyTrpPheGlyGluGlnPheAlaGlnMet 340
 Db 961 CACATCAAACTGATCGGTCAACCTGGTAAATGTTGCGCGACGCAATTCGCGCAGATG 1020
 Oy 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyThrAlaArgAlaSerAlaSer 360
 Db 1021 ATAGGCGCGAAACCCCTGTGACAAATAGGGATCTTGCCTGCTTCTGCTTC 1080
 Oy 361 AsnValAspAspMetCysLeuIleLysSerCysAlaAspLeuAlaGluCysAlaPhe 380
 Db 1081 AACGTTGACGACATGGCTTGTGATCAATCTGCGCGCATTTGGCGGTCCAGTCCGCTTC 1140
 Oy 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
 Db 1141 CCGCGGAGTCTGGCGTATCGGTCAAGCAAGACACGCGCAACGTTGCGTGGATC 1200
 Oy 401 GIupheProxArgIlelysgIylybProphesAnIleasPhrThraPheAnSer 420
 Db 1201 GAGTTTCGCGCATCAAGGACGGCAACCGTTCAATATGACACCACTGGTTCAATAGC 1260
 Oy 421 MetLeuSerGluIleGIyGlnProlysgIylybValGluValSerHis 437
 Db 1261 ATGTTGAGGAATCGGCAAGCCTTAAGCGGTAAGTGAAGTCAAGTCAAGCAC 1311
 RESULT 6
 AX398428 1311 bp DNA linear PAT 27-MAY-2002
 LOCUS Sequence 5 from Patent WO0220796.
 DEFINITION AX398428
 ACCESSION AX398428
 VERSION AX398428.1 GI:21261167
 KEYWORDS
 SOURCE
 ORGANISM
 Methylobionas sp.
 Bacteria; Proteobacteria; Gammaproteobacteria; Methylcoccales;
 Methylcocccaceae; Methylobionas.
 REFERENCE
 1 Koffas,M., Odum,J.M. and Schenzie,A.
 High growth methanotrophic bacterial strain Methylobionas 16a

JOURNAL Patent: NO 0230728-A 5 14-MAR-2002;

E.I. DUPONT DE MEMOIRS AND COMPANY, Legal Patent Records Center

FEATURES
source Location/Qualifiers
1.1311
/organism="Methylobacter sp."
/mol_type="unassigned DNA"
/db_xref="taxon:418"

ORIGIN

Alignment Scores:
Pred. No.: 2.12e-161 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-701-200-6 (1-437) x AX398428 (1-1311)

1 AapValValThrTyrProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhePhe 20
1 GATGTGGTGCACATGCGCCCTATCATTAACGGCTGATATTCGATTGTCATGTTT 60
21 LeuAsnPheAsnPheTyrThrLeuMetAsnLysProLysLeuValAlaIleLeuThrAla 40
61 CTTAACTTAACTTCTACACGCTCATGAACAACTTAAAGTTGCAATGCTGACAGCA 120
41 G1yG1yLeuAlaProCysLeuAsnSerAlaIleG1ySerLeuIleG1yUarG1yTrgIu 60
121 GCGCGCTTGGCCCTGTTTGAATTCGCAATCGGTAAGTTGATCGAACGTTATACGAA 180
61 IleAsnProSerIleG1yIleIleCysTyrArgIyG1yTyrLysG1yLeuLeuLeuG1y 80
181 ATGATCTCAGCATAGAAATCATTTGCTATCCGGCGGTTATAAGCCCTGTTGCTGGGC 240
81 AapSerTyrProValThrAlaGluValArgLysLysAlaG1yValLeuGlnArapPheG1y 100
241 GATTTCTTATTCAGTAACGGCCGCAAGGCGTAAAGGCGGCGTCTGCAACGTTTGGC 300
101 G1ySerValIleG1yAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
301 GGTTCGTGATCGGCAACGCGCGCTCAATGACCAATGTCAAAGACTGCGTGAAGCC 360
121 G1yLeuValLysG1yG1yG1yAspProG1yLysValAlaAlaAspG1yLeuValLysAsp 140
361 GGTTCGTCAAGAGGCGTGAAGATCCGCAAAAGTCGCGCTGATCAATGGTTAAGAT 420
141 G1yValAspIleLeuHisThrIleG1yG1yAspAspThrAsnThrAlaAlaAspLeu 160
421 GGTTCGATATTCGACACACATCGCGCGGCGATGATACCAATACGCGACGCGGATTTG 480
161 AlaAlaPheLeuAlaArgAsnAsnTyrG1yLeuThrValIleG1yLeuProLysThrVal 180
481 CGAGCATTCCTGGCCGAAATTAATACGACCTGACCGTCATTTGTTACTTAAACCGTC 540
181 AapAsnAspValPheProIleLysG1ySerLeuG1yAlaATrPThrAlaAlaG1yGlnG1y 200
541 GATACGACGATATTCCTGATCAAGCAATCATAGGTTGTTGACGCGCGCAAGCAAGGC 600
201 AlaArgTyrPheMetAsnValValAlaG1yAsnAsnAlaAsnProArgMetLeuIleVal 220
601 GCGCGTTATTTCTATGAACGTGGTGGCGCAAAACCAACGCGCAACGCGATGTCGTA 660
221 HisIleValIleMetG1yArgAsnCysG1yTrpLeuThrAlaAlaIleThrAlaG1yGlnIle 240
661 CACGAGTGTATGGCGCTTAATCGCGCTGCTGACCGCTGCAACGCGCGAGAAATATCC 720
241 LysLeuLeuAspArgAlaG1yTrpLeuProG1yLeuG1yLeuThrArgIleUarG1yTrgIu 260
721 AAATTAATGACCGCGAGTGGTGGCGGAAATGGGTTTGAATCGTGAATCTTATGAA 780
261 ValHisAlaValPheValProG1yMetAlaIleAspLeuG1yAlaG1yAlaValArgLeu 280

781 GTGACGCGGATTCGTTCCGGAATGCGATCGACTCGAAGCGGAAGCCAAAGCCCTTG 840
281 ArgGluValMetAspLysValAspCysValAlaAsnIlePheValSerG1yG1yAlaG1yVal 300
841 CCGCAAGTGTATGCAAAAGTGTATGCTGATGCTGCAACATCTTCGTTGCAAGGTCCGCGTC 900
301 G1yAlaIleValAlaG1yMetGlnAlaLysG1yGlnG1yValProArgAspAlaPheG1y 320
901 GAAGCTATCGTCGGGAAATCAGGCCCAAGGCCGGAAGTCCCGCGGATGCTTCGCGC 960
321 HisIleLysLeuAspAlaValAsnProG1yLysTrpPheG1yG1yGlnIlePheAlaGlnMet 340
961 CACATCAACTGATGCGGTGCAACCTGTAATGCTTCGCGAGCAATTCGCGAGATG 1020
341 IleG1yAlaG1yLysThrLeuValGlnLysSerG1yTyrPheAlaArgAlaSerAlaSer 360
1021 ATAGCGCGGAAAAACCCGTGACAAATAATCGGGAATCTTCGCGCGTCTTCCTCC 1080
361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValG1yCysAlaPhe 380
1081 AACGTTCAGCATGCGTTGATCAATCGTCCGCGACTTGGCGGTGAGTGCAGCTTC 1140
381 ArgArgG1ySerG1yValIleG1yHisAspGluAspAsnG1yAsnValLeuArgAlaIle 400
401 GluPheProArgIleLysG1yG1yLysPheAsnIleAspThrAspTrpPheAsnSer 420
1201 GAGTTCCGCCCATCAAGCGCGGCAACGTTCAATATCGACACGCACTGTTCAATAGC 1260
421 MetLeuSerG1yIleG1yGlnProLysG1yLysValG1yValSerHis 437
1261 ATGTTGACGAAATCGGCCACCTTAAAGCGTAAAGTCGAAGTCAGCGCAC 1311

RESULT 7
LOCUS BX294151/c 295650 bp DNA linear BCT 11-JUL-2003
DEFINITION Pirellula sp. strain 1 complete genome, segment 19/24.
ACCESSION BX294151 BX119912
VERSION BX294151.1 GI:3246812
KEYWORDS Rhodopirellula baltica SH 1
SOURCE Rhodopirellula baltica SH 1
ORGANISM Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
Planctomycetaceae; Rhodopirellula.
REFERENCE 1 (bases 1 to 295650)
Gloeckner, F.O., Kube, M., Bauer, M., Teeling, H., Lombardot, T.,
Ludwig, W., Gade, D., Beck, A., Borzym, K., Heilmann, K., Rabus, R.,
Schlesner, H., Amann, R. and Reinhardt, R.
Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1
Proc. Natl. Acad. Sci. U.S.A. 100 (14), 8298-8303 (2003)
JOURNAL MEDLINE 22735913
PUBMED 12835416
2 (bases 1 to 295650)
Kube, M., Borzym, K., Heilmann, K., Klages, S., Marguardt, I.,
Lehrack, S., Beck, A., Pawlik, R., Reinhardt, R., Gloeckner, F.O.,
Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Rabus, R.,
Schlesner, H. and Amann, R.
Direct Submission
Submitted (21-JAN-2003) Max Planck Institute for Molecular
Genetics, Potsdammer Platz 1, D-10585 Berlin, Germany
Planck Institute for Marine Microbiology Celsusstrasse 1, D-20359
Bremen, Germany
This project was carried out by
*Max Planck Institute for Molecular Genetics, Berlin, Germany; *Max
Planck Institute for Marine Microbiology, Bremen, Germany; in the
framework of the REXG-Project, <http://www.rexg.de> -----
Genome Center
Center: Max Planck Institute for Molecular Genetics
----- Summary Statistics -----

Sequencing vector: pUC19; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 7142861 bases at least Q40
Consensus quality: 7145318 bases at least Q30
Consensus quality: 7145484 bases at least Q20
Quality coverage: 8.03

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid sequence; assembly was additionally confirmed by long range PCR and cosmid end sequences.

See <http://www.micro-genomes.mpg.de/pirellula/> for more information on including minimal tiling path from a set of 220 cosmids out of 908. See the `misc_feature` tag below for the boundaries of the MTP cosmids. ----- Annotation

Center: Max Planck Institute for Marine Microbiology
 Celsiusstrasse 1, D-28359 Bremen, Germany.

Center Code: MPIM
Email: fog@mpi-bremen.de
Phone: +49 (0)421 2028 938 Fax: +49 (0)421 2028 580

Three different programs (Glimmer, Critica and Orpheus) were used for ORF-prediction. A nonredundant list of ORFs was generated by suitable parsing of the results. Automated annotation was done with the software package Pedant Pro (http://www.biomax.de). All ORF predictions and annotations were manually corrected by considering all results of the different tools applied. See <http://www.regg.de> for more information and access to supplementary information.

FEATURES

Location/Qualifiers

1. .295650

```
/organism="Rhodospirillum rubrum SH 1"
```

/mol_type="genomic DNA"

```
/strain="1"
```

/db_xref="taxon:243090"

```
complement (414. .1508)
```


CDS

/gene="Pipifk"
48..1280
/gene="Pipifk"
/EC number="2.7.1.11"
/function="converts fructose-6-phosphate to
fructose-1,6-bisphosphate"
/note="Involved in glycolysis"
/codon_start=1
/product="pfi-phosphofructokinase"
/protein_id="AAF70463.1"
/db_xref="GI:7862073"
/translation="MKTVALLTAGGLAPCLSSANGGLIERTTMSPEPTNILLNGY
KGLLSEKLVTPARKLQHVLTGGSCIGNSRYMANAVACVKRGLVEGQDPKQV
AADQLIKGVDTLTIGDPTVAADLAALYKAHGYTLRVLGPTINDIVPVOS
LGAMTAAEQASRFQNVVAEQTANPVLVHEVMGSGCYLTAQADYYRAOLAHF
AMELHETREYRDIHAVYPERMTIDLKABARLRAVMEYGVNFIISBEGINDIYAE
MTAKETVPRDPFGVHKIDLTIPGAMFGQFGMVGADVLYVOKSYFERSRANPANE
LRLTGMVDLAVDCAIRGERAGLIGHDEERNVLRALIFERVGAFAFNIDHFWFTLL
NEIGOPKGAKVSAHAGDE"

ORIGIN

Alignment Scores:

Pred. No.: 2,66e-99 Length: 1427
Score: 1453.00 Matches: 272
Percent Similarity: 80.74% Conservative: 55
Best Local Similarity: 67.16% Mismatches: 78
Query Match: 63.62% Indels: 0
DB: 3 Gaps: 0

US-10-701-200-6 (1-437) x AFA246209 (1-1427)

QY 33 LysLysValAlaIleLeuThrAlaGlyLysLeuAlaProCysLeuAsnSerAlaIleGly 52
DB 54 AAGACGTGGCCCTGCTCACTGCGGAGGCGCTGCCCTGCTGCTGCTGCGCGGCG 113
QY 53 SerLeuIleGluArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGly 72
DB 114 GGCCTCATGACGCGCTACACCGAGATGCGCCGAGAACCAACATCATCTTCACTCAAC 173
QY 73 GlyTyrLysGlyLeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLys 92
DB 174 GGCCTCAAGGGCGCTGCTGCGGAGAGAGCTCTGCTAACCCTGCGCGCGCGCTCCAG 233
QY 93 AlaGlyValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThr 112
DB 234 GCGCACTGCTCTGCAACCTCGCGGCACTGCTGCGCACTCGCGGCTCAAGATGCGC 293
QY 113 AsnValLysAspCysValLysArgGlyLeuValLysGlyGluGluAspProGlnLysVal 132
DB 294 AACGTGCGCGACTGCTCAAGCGCGGCGCTGCTCAAGAGAGGCGCAGGACCGCGAGGTG 353
QY 133 AlaAlaAspGlnLeuValLysAspGlyValAspIleLeuHisThrIleGlyGlyAspAsp 152
DB 354 GCGGCGGACCTGCTCAAGAGCGCGGTGAGCTGCTCAACCGATCGCGCGCGCGAGAC 413
QY 153 ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuValArgAsnAsnTyrGlyLeuThr 172
DB 414 ACCAACAACGCGCGCGGCGGACCTGCGGCGGTACTCAAGCGCGGCGGTACACGCTGCC 473
QY 173 ValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSerLeuGly 192
DB 474 GTGATCGGCTGCGCCCAAGCATGACAGCATGCTGCGCGCGCGCGCGGTGCGGCG 533
QY 193 AlaTyrPheAlaAlaGluGlnGlyAlaArgTyrPheMetAsnValValAlaGluAsnAsn 212
DB 534 GGCATGACGCGCGCGCGCGGCTCTTCTTCAAGAACGTCGTCGCGGAGCGAGCG 593
QY 213 AlaAsnProArgMetLeuIleValHisGluValMetGlyArgAsnCysGlyTyrPheLeuThr 232
DB 594 GCGAACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653
QY 233 AlaAlaThrAlaGlnGlnTyrArgLysLeuLeuAspArgAlaGluTyrPheProGlnLeu 252

DB 654 GCCCAGCGCGCGGACTACTACCGCGCGGCGAGCTTCGCGCACCGCGGATTCGCGCGGAGCTTC 713
QY 253 GlyLeuThrArgGlySerTyrGluValHisAlaValPheValProGluMetAlaIleAsp 272
DB 714 GCGCACACCGCGGAGGCTTACGATCCAGCGGCTTACGTCGCCGAGATGACCATGCGAC 773
QY 273 LeuGluAlaGluAlaLysArgLeuArgGluValMetAspLysValAspCysValAsnIle 292
DB 774 CTCAGCGCGGAGCGCGCGCGCTGCGCGCGCTCATGAGCGCGGTGCGGTGCGTCAACATC 833
QY 293 PheValSerGluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAlaLysGln 312
DB 834 TTCCTCAGGAGCGCGCGGCGGATCAACGATCTGTCGCGAGTTGACCGCGAAGCGCGAG 893
QY 313 GluValProArgAspAlaPheGlyHisIleLysLeuAspAlaValAsnProGlyLysTyr 332
DB 894 ACCGTCCCGCGGACCCCTTCGCGCGCGATGATGATGATGATGATGATGATGATGATG 953
QY 333 PheGlyGluGlnPheAlaGlnMetIleGlyAlaGluLysThrLeuValGlnLysSerGly 352
DB 954 TTCGCAACAGTTTCGCGCGGATGTCGCGCGGCGGACAGGTCCTGTCGGAAGTCGCGG 1013
QY 353 TyrPheAlaArgAlaSerAlaSerAsnValAspAspMetArgLeuIleLysSerCysAla 372
DB 1014 TACTTCTCCGCTGCG 1073
QY 373 AspLeuAlaValGluCysAlaPheArgArgLysSerGlyValIleGlyHisAspGluAsp 392
DB 1074 GACCTGCGCGCTGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1133
QY 393 AsnGlyAsnValLeuArgAlaIleGluPheProArgIleLysGlyGlyLysPheAsn 412
DB 1134 CGCAACGCGGTGCTGCGCGCGATCGATTGAGCGGTCAAGGCGCGCGCGCGCGCGTCAAC 1193
QY 413 IleAspThrAspTyrPheAsnSerMetLeuSerGluIleGlyGlnProLysGlyLys 432
DB 1194 ATCGACACCGCGGTTCACCACTCTCAACGAGATGCGCGCGCGCGCGCGCGCGCG 1253
QY 433 ValGluValSerHis 437
DB 1254 GTCTCGGTGCGCGC 1268

RESULT 9
LOCUS CQ363740 8530 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 23 from Patent WO0181581.
ACCESSION CQ363740
VERSION CQ363740.1 GI:41300434
KEYWORDS
SOURCE
ORGANISM
Propionibacterium acnes
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Propionibacteriaceae; Propionibacterium.
REFERENCE
AUTHORS
1 Skelky, Y.A., Persing, D.H., Mitcham, J.L., Wang, S.S., Bhatia, A.,
L'Homme, J.F., Zhang, Y., Jen, S., and Carter, D.
TITLES
Compositions and methods for the therapy and diagnosis of acne
vulgaris
JOURNAL
Patent: WO 0181581-A 23 01-NOV-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
1..8530
/organism="Propionibacterium acnes"
/mol_type="unassigned DNA"
/db_xref="taxon:1747"

ORIGIN

Alignment Scores:
Pred. No.: 9.93e-98 Length: 8530
Score: 1444.50 Matches: 274
Percent Similarity: 80.59% Conservative: 54
Best Local Similarity: 67.32% Mismatches: 74
Query Match: 63.24% Indels: 5

[illegible]

| | | | | | | | | | | | | | |
|---|--|---|-----|--|--|--|--|--|--|--|--|--|--|
| ACCESSION | 1-phosphotransferase (PFK) gene, complete cds. | | | | | | | | | | | | |
| KEYWORD | M67447 | | | | | | | | | | | | |
| KEYWORDS | pyrophosphate-fructose 6-phosphate 1-phosphotransferase. | | | | | | | | | | | | |
| SOURCE | Protonibacterium freudenreichii | | | | | | | | | | | | |
| ORGANISM | Protonibacterium freudenreichii Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Protonibacteriaceae; Protonibacteriaceae; Protonibacterium. | | | | | | | | | | | | |
| REFERENCE | 1 (bases 1 to 1400) | | | | | | | | | | | | |
| AUTHORS | Ladoro, U.S., Gollapudi, L., Tripathi, R.L., Latchaw, S.P. and Kemp, R.G. | | | | | | | | | | | | |
| TITLE | Cloning, sequencing, and expression of pyrophosphate-dependent phosphofructokinase from Protonibacterium freudenreichii | | | | | | | | | | | | |
| JOURNAL | J. Biol. Chem. 266 (25), 16550-16555 (1991) | | | | | | | | | | | | |
| MEDLINE | 91358443 | | | | | | | | | | | | |
| PUBMED | 1653240 | | | | | | | | | | | | |
| COMMENT | source text: P.freudenreichii DNA. | | | | | | | | | | | | |
| FEATURES | Location/Qualifiers | | | | | | | | | | | | |
| source | 1..1400 | | | | | | | | | | | | |
| gene | /organism="Protonibacterium freudenreichii" | | | | | | | | | | | | |
| gene | /mol_type="genomic DNA" | | | | | | | | | | | | |
| gene | /db_xref="taxon:1744" | | | | | | | | | | | | |
| gene | 74..1288 | | | | | | | | | | | | |
| gene | /gene="PFK" | | | | | | | | | | | | |
| gene | 74..1288 | | | | | | | | | | | | |
| gene | /gene="PFK" | | | | | | | | | | | | |
| gene | /BCD_number="2.7.1.90" | | | | | | | | | | | | |
| gene | /codon_start=1 | | | | | | | | | | | | |
| gene | /transl_table=1 | | | | | | | | | | | | |
| gene | /product="pyrophosphate-fructose 6-phosphate 1-phosphotransferase" | | | | | | | | | | | | |
| gene | /protein_id="AA25675.1" | | | | | | | | | | | | |
| gene | /db_xref="GI:150931" | | | | | | | | | | | | |
| gene | /translation="MVKKVALITNAGPAGLSSALIELIKRYEVSPEITLIGRYGKGLKDGSLSPAVRAHYDRILRSFGSPPLGNSATKLTWKDYLRKGLVASGDPLPKYAAQDLADGVADVLTIGDDTNTTAAALAYLAQHDPLTVVGLPKTIDNDVIRQSLGAMTADDEGAFANVYPAATHIDLAEARLRTVMDEGVSVNFISEGAVPIVAGLPEAGDRLRGMDIHAYLPBATDIDLAEARLRTVMDEGVSVNFISEGAVPIVAGMOATGSEVPTDAFGHVLQDKINPAMFAKQFAEIRIGAKTMVOKSGFSRKAQNAQLEIATATMAVADALAGTGGVGVQDEADQKLSVIDPKRIAGHKPDITLDWYTGOLLARIGQAPAPIAA" | | | | | | | | | | | | |
| ORIGIN | | | | | | | | | | | | | |
| Alignment Scores: | | | | | | | | | | | | | |
| Pred. No.: | 6.23e-89 | | | | | | | | | | | | |
| Score: | 1314.00 | | | | | | | | | | | | |
| Best Local Similarity: | 77.78% | | | | | | | | | | | | |
| Best Local Similarity: | 62.63% | | | | | | | | | | | | |
| Query Match: | 57.53% | | | | | | | | | | | | |
| DB: | 1 | | | | | | | | | | | | |
| DB: | Gaps: 0 | | | | | | | | | | | | |
| US-10-701-200-6 (1-437) x PRSPFK (1-1400) | | | | | | | | | | | | | |
| OY | 33 | LyelYbValAlaleLeuthrAlGlyGlyLeuAlaProCYbLeuAenSerAlalleGly | 52 | | | | | | | | | | |
| Db | 80 | AAAAAGGCGCTCGCTGACCGCGTGGCTTC | 139 | | | | | | | | | | |
| OY | 53 | SerleuilegIuAqYrYrthGluIlaAspSerleuilegIuileIcYsYrAqGly | 72 | | | | | | | | | | |
| Db | 140 | GAGCTCATCAAGCCGTATACCGAGGTATCACCCGAAACGACCTTCATGCGTAT | 199 | | | | | | | | | | |
| OY | 73 | GlyYrYrYbGlyLeuLeuLeuGlyYbAspSerYrProValThralaGluValaYrYbYs | 92 | | | | | | | | | | |
| Db | 200 | GGCTATGAGGCGCTGCTCAAGGGGATTCCTCGAGTTTCCTCGCGTGGCGACAC | 259 | | | | | | | | | | |
| OY | 93 | AlaGlyValLeuGluAqPheGlyYsValIleGlyAenSerAqValYbLeuThr | 112 | | | | | | | | | | |
| Db | 260 | TACGACCGGCTTCAGCTTCGGCGGTATCCGATCGGAACTCCGGGTACAGCTACC | 319 | | | | | | | | | | |
| OY | 113 | AsnValYbAspCYbValYbAqGlyLeuValYbGluGluYbAspProGluAlaYbVal | 132 | | | | | | | | | | |
| Db | 320 | AAATGTAAGACCTGGTTGGCGGGGCTGCTTCCTCCGCAATGATCCCTCAAGTT | 379 | | | | | | | | | | |

| | |
|-----------|---|
| TITLE | Boillard,P., Becker,A., Bouvier,M., Cadet,E., Dreaño,S., Gloux,S., Godié,T., Goffeau,A., Kahn,D., Kias,B., Lelubre,V., Masny,D., Pohl,T., Portetcelle,D., Puehler,A., Purnelle,B., Ramsperger,U., Renard,C., Theault,P., Vandemol,M., Weidner,S. and Gallibert,F. Analysis of the chromosome sequence of the Legume symbiont Sinorhizobium meliloti strain 1021 |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001) |
| MEDLINE | 21396507 |
| PUBMED | 11481430 |
| REFERENCE | 2 (bases 1 to 294800) |
| AUTHORS | Gouzy,J. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO EU Consortium |
| COMMENT | MELILO EU Consortium: Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, UMR215-CNRS-INRA, BP277, F-31326 Castelnau, France, Laboratoire de Génétique et Développement UMR6061-CNRS, Faculté de Médecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes, France, GARC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany, Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25, D-33615 Bielefeld, Germany, Unité de Biochimie physiologique, Université Catholique de Louvain, Place Croix du Sud 2, Bte 20, B-1348 Louvain-la-Neuve, Belgium, Unité de Microbiologie, Faculté des Sciences Agronomiques de Gembloux, Avenue Mancelain Jura 6, B-5030 Gembloux, Belgium. E-mail:derome.gouzy@ceoulouse.inra.fr http://sequence.colouence.inra.fr/meliloti.html . Location/Qualifiers |
| FEATURES | source |
| gene | 1. 294800 /organism="Sinorhizobium meliloti" /mol_type="Genomic DNA" /strain="1021" /db_xref="taxon:382" complement(147..1382) /gene="SMC04259" complement(147..1382) /gene="SMC04259" complement(147..1382) /function="cell processes: transport of small molecules" /note="Product confidence : putative Gene name confidence : hypothetical predicted by Codon usage predicted by Homology predicted by Framed" /codon_start=1 /evidence=not_experimental /transl_table=11 /product="PUTATIVE PERIPLASMIC BINDING ABC TRANSPORTER SIGNAL PEPTIDE PROTEIN" /protein_id="CAC46508.1" /db_xref="GI:15074951" /db_xref="GDB:O92P64" /db_xref="UniProt/TREMBL:O92P64" /translation="NMLRLAALAGTALPRGASATDLVTHWTSGGGAAVAELAAAPATGKNWDAGAGSGGTARPMIRITGDDPMATQFNHGQAEELVQAGLMRLDITDAKTENKKEIVPSLSIDCTEGEIVGAPVNIHSOMLNLNAAFKAGVEVPRKMDDEVNAPALEKAGIVPLVAGCGPOMANGAFVLAIAKGKPEKVFQKDEVAAGPELAKFYKAADARRKSKGTNVODNQTNTVITGKAGGCIKMDKAGGQALQAKGVDITCLPGIGVNEVISTGDAPLFFLLEDESKAGDVALSTLLKPTQVAFNIKKSGLPRGVDVLAANDCMKRGDIILAKGNVIGCTDQLLSADSQKEDLFSEFFANH |
| CDS | |


```

gene
CDS
    /product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"
    /protein_id="CAC46509.1"
    /db_xref="GI:15074952"
    /db_xref="GOA:Q92P63"
    /db_xref="UniProt/TREMBL:Q92P63"
    /translation="MTGMDGKTRNRTTAAPEGGPPLTKTIAFMTGLGVTVSRALK
    DAPDIAETKERVRLVAKQIGYQPNBAGVRLRTGKTNVLSLVLTLEEIMGTSPNVI
    GTEILAGTQYHLVTVPSASADPLGPRIYIILDTGADGVIISRTSPNDPRTVLTTER
    RLFPATHGRTEMGLIHPYDFONERPAYEAVRLVDRGRRLVLEPPNLPFHSMR
    TGEGRVDRDGAEAVSEFHQVNIIDHSLVAIRDAEKLMRSPDAPGVSGSGGAILI
    ADEAKGKVEDAEVSKVPSDFLRMLRPEVMTMEDIRIAGRELAKAVIGHIEGR
    PETLSQSPERFQPMARSASVS"
    /length=3023
    /note="Sm-4 OR SMC04506
    REPEAT SM-4
    predicted by Homology"
    /evidence=not experimental
    /complement(3036..3302)
    /gene="SMC04261"
    /complement(3036..3302)
    /function="miscellaneous; unknown"
    /note="Product confidence : hypothetical
    Gene name confidence : hypothetical
    predicted by Codon usage
    predicted by FrameD"
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="HYPOTHEICAL PROTEIN"
    /protein_id="CAC46510.1"
    /db_xref="GI:15074953"
    /db_xref="UniProt/TREMBL:Q92P62"
    /translation="MKETVIEGSKAEVTAASVTIRDAERRIRKAGYEQRVRSLSA
    TCAPRIGPIQYLMQIDFVAGKLEQLNPISPTDIXKYPAGI"
    /complement(3407..4837)
    /gene="gnd OR SMC04262"
    /complement(3407..4837)
    /gene="gnd OR SMC04262"
    /BC_number="1.1.1.44"
    /function="small molecule metabolism; energy metabolism,
    carbon; oxidative branch, pentose pathway"
    /note="Product confidence : probable
    Gene name confidence : putative
    predicted by Homology
    predicted by FrameD"
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="PROBABLE 6-PHOSPHOGLUCONATE DEHYDROGENASE
    (DECARBOXYLATING) PROTEIN"
    /protein_id="CAC46511.1"
    /db_xref="GI:15074954"
    /db_xref="GOA:Q92P61"
    /db_xref="UniProt/TREMBL:Q92P61"
    /translation="MSQAEIGILIGVMSNLSLANTAEKGNRIAVENRTVDAIRKRYA
    EAGALKDQIVPCETIEEYFAIRPPRIITIMKAGSPVUQNEALKPHLAKDINIDA
    GNAERDITRRFDALDQSGITFGKGVSEEGARHPSIMWGTESEIRRVKVLTS
    IAAKYDSDPCVAMLGNGAHFKYTIHNSIEVADQMIAEIVGILGDKLMAOEIGE
    VEGAMNKGRLNSYLEITERKVLKAADPLGKPIVDMITDKAGQKGGKSVLEAOMG
    VVATAIEAAVVAARSISAKEEREAEKVLGPVGIKYADRGGFKDLENLLAKKI
    GAYAGFAVMAASKEFGNMLPMPITAKIPRAGCIRSGFLDEITTAFTKADPAANLI
    VPPAPAAKESDGAIRRVSTAVGLGVYPALASLGYFDSTRRGRTANYTQAORD
    FFGAHGFRVDGDSHHGPMGSLNA"
    /length=5273
    /gene="SMC04263"
    /length=5273
    /gene="SMC04263"
    /function="cell processes; transport of small molecules;
    amino acids, amines, peptides
    /note="Product confidence : putative
    Gene name confidence : hypothetical

```

```

gene
CDS
    /product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"
    /protein_id="CAC46512.1"
    /db_xref="GI:15074955"
    /db_xref="GOA:Q92P60"
    /db_xref="UniProt/TREMBL:Q92P60"
    /translation="MDTIIGFNTLIFMGVYLTVGLAVGLYFTVRIGFPOIVSEME
    RYLSRGGSDAAGISPPQALMYSRVGTGNLAGVAVLVLGGPGATFMMIVAVFG
    MATAVESALAOYLKIRNEDGOYRGGFYIRGLAVALPAWATIFSACILISFGLVFA
    VANSIADAVOGAFGVKAVGAVVALSLGSLVITFGIIRQIARAEIVVPMMAAYLT
    AYVULANASAVPULMTISSAFGQEAAGITGCIANAANGVIRGGLFSPNAGGS
    AVNIAVATVPVPHBSSQGFVOSLGVFTITLICTATSVWILSLGTLFSGSVTGTOL
    TQAMSTHIGSAGTYFIAIAIFFAFSTIIIGNSVAMNLTVYGGNRLGLMTMRCAT
    LAMVWGAVESITVFDADADASMGLMATINLVALILSGTIAKLTROYFEQRKAGAVP
    VFHADYPELOKIDGIEINSRD"
    /length=6773
    /gene="SMC04264"
    /length=6773
    /gene="SMC04264"
    /function="miscellaneous; not classified regulator"
    /note="Product confidence : putative
    Gene name confidence : hypothetical
    predicted by Codon usage
    predicted by Homology
    predicted by FrameD"
    /codon_start=1
    /evidence=not experimental
    /transl_table=11
    /product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"
    /protein_id="CAC46513.1"
    /db_xref="GI:15074956"
    /db_xref="GOA:Q92P59"
    /db_xref="UniProt/TREMBL:Q92P59"
    /translation="MGTSCPAYDKDARARERLOQEARLRLGPALAQTSIVPEQSP
    ALATRFKCLMHRNPDGAPSSIAVVPDCCSIDINSSKALAVGPRIAPFLPDLGE
    TIGAFRIGAAASWLRPLREMTGTATVLAHLVWGAVTEIERIRDAISTERDLVL
    AALIGRLSAABPPPDIAACVAHLOVGSITVDAIRTLVREIGTSERTLRRCBHL

```

Alignment Scores:

| Pred. No.: | Score: | Length: | Matches: | Mismatches: | Indels: | Gaps: |
|------------------------|---------|---------|----------|-------------|---------|-------|
| 3,12e-85 | 1302.00 | 294800 | 246 | 59 | 90 | 0 |
| Percent Similarity: | 77.22% | | | | | |
| Best Local Similarity: | 62.28% | | | | | |
| Query Match: | 57.01% | | | | | |

```

US-10-701-200-6 (1-437) x SME591789 (1-294800)
QY 33 LysLysValAlaIleLeuThrAlaGlyGlyLeuAlaProCyLeuAunserAlaIleGly 52
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 273560 AAGAAAGTCGCAATGTGACGCGAGCGGCTCGCCCTTCATCGCTGTCGCG 273619
QY 53 SerLeuIleGluArgTyThrGluIleAapProSerIleGluIleIleCyStyTArgGly 72
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 273620 GGCCTGATCGAAGCGTACCGCAGTCGCGCGAGCTAGAGCTCGTGCCTCC 273679
QY 73 GlyTyrTyLeuGlyLeuLeuLeuGlyAapSerTyTyrProValThrAlaGluValAlybly 92
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 273680 GGCCTATCAGAGCGCTGCTCTCGCCGACCGAGTCGAATATACCCCGGCATGCGGAGAG 273739
QY 93 AlaGlyValLeuGlnArgPheGlyGlySerValIleGlyAunserArgVallyblyLeuTnr 112
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 273740 GCGCATGTCTCCACCGCAGCGGATGCGCATGCGGACACGCGCGTGAAGCTCAC 273799
QY 113 AenVallyAapCyAlyVallyAargGlyLeuVallyblyGlyGlyGluAapProGlnlyVal 132
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 273800 AACACTGCCGACTGCTGGAAGCGCGCTCGTCACGAAGAGGGAGAAATCCTTTGCCGCTC 273859

```



```

gene
3439..4068
/ gene="AGR_C_3823"
/ protein_id="AAK87857.1"
/ db_xref="GI:15157243"
/ note="HYPOTHETICAL 45.0 KD PROTEIN IN TOLC-RIBB INTERGENIC REGION (ORFA) (0386)"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3823p"
/ protein_id="AAK87857.1"
/ db_xref="GI:15157243"
/ translation="MSDSAECCGTTIYLMDCAIQAGHRAQFLDIRDIGDASGRYADL OSRVIERCKLYPMERPELRESPANLPASGPIFIRAPKALLSNGLPMLMORRNH PNLPSFPDPDEITELQDYVYKPLSRGENVSLFRNRREILSPAGSTGKGYVQQA YAPLPSERGVAVLSWIVGDRSCGLGREVDVSPITANLSRYVPHIEG"
/ complement(4553..5794)
/ gene="AGR_C_3826"
/ complement(4553..5794)
/ note="SARCOSINE OXIDASE"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3826p"
/ protein_id="AAK87858.1"
/ db_xref="GI:15157244"
/ translation="MQATRMKRVNGTCKEGSGSHFNITYVGRGMGAANAHLAETV DGVALLIGEPADIKSHQGVFASHYDEARITTTIGDDPMALLANRSIARYADIFARS GVEFYAPVCLMVGPERGANGPFDVLAARLIGVSTELLGDOSLKSFPYFSEEPG CEVPEPNDAGVGNPRALVKAQALAEKAVTLDIIVYSTRDEGRASVOTAAGVY TAERYLVAAAGSFTIRDLIPQVGLDYARTVAFPEIDRGLGOYAGMPSLISEYHDT TKHTYLLPVRYPDQKFTYIKGCGDDPDRKVGSDPEIRERFSGSGRESVRDHSYGT LVPYDHSKVSMAACVSKTKSGYPALIGFTASPRIAIVLGGNGTAKSSDEIGRLGAA LIRDKIADDASTDQKPEFL"
/ complement(5803..7116)
/ gene="AGR_C_3830"
/ complement(5803..7116)
/ note="Chain A, Complex Of Monomeric Sarcosine Oxidase With The Inhibitor Dimethylglycine"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3830p"
/ protein_id="AAK87859.1"
/ db_xref="GI:15157245"
/ translation="MROAGIAGELRPASIIICEIAKNDGSGRAHVAVNRKSDTEHDMP DEXYLVGGMGAARHLISYOTGVALLIGAPADRKTHQGVSHSVDRVIRRGF DGDPMWALALRSIRRYAIEAKSGIRFTEAGCLFTGNGKLAGDVPYASDRLL GLGVETGAELAGRPMPALPADGSGSEANACHINRAIVKQCAIAEAGCGRVL RETAAHIRTGSSEVMTREGATYTAEKYVAGGRTNMAALPSPVDAATAGRTIVF FELDARQALFSAMPSTIVLAETEDIVILPVRYPDKVYLKIGSESEKRLSTLV QAVDFHSDGTDEVEFLTKKALSLMPALAGCPVSGSCVASTGSGYIITYTOSIN IAVLTGCVNVAKSSDEIGRLGALLNQLTEDEFAEMSVFV"
/ complement(7083..7670)
/ gene="AGR_C_3831"
/ complement(7083..7670)
/ note="putative transglycosylase"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3831p"
/ protein_id="AAK87860.1"
/ db_xref="GI:15157246"
/ translation="MPFHCCHPIWRILFLMNRPVVAVAVGVALASQAGISPADED

```

```

gene
VKTQVLTLETFRKQGPPIPERKSLPASLKNYSDLIVTKAKRYGVPTNLAAHVSVES KENPARAGSAGVAMQIKPATARYMGEFGATKALYDPPTNIRMGQYLATAHQLGGG EVCSITILRNAGAGATRMNVPKRCVKQALLAS"
/ complement(7870..8649)
/ gene="AGR_C_3833"
/ complement(7870..8649)
/ note="conserved hypothetical protein PA5485 (imported) - Pseudomonas aeruginosa (strain PAO1)"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3833p"
/ protein_id="AAK87861.1"
/ db_xref="GI:15157247"
/ translation="MKECIPDPADPFGCAVAPSPNHGRLGVAGPDIIILLHYTGWT ADGALSWLRNPSQVSSHVEFEDRVIQOLVPSRANMAGSAGDEIDNSRIGI EIANQCHRGGLPEPEAOVAAVIELCRDGRWSTAPRVLASHVAPLRKYDPEKFI PNDLSQHGCHWVEPAPIRGRFFROGDHQPVEALQSMSTIYTGAEITGAVCCKT EGAVAAPQHFPRPALVDGIADFTITDILRLISLPRFSAA"
/ complement(8646..9422)
/ gene="AGR_C_3834"
/ complement(8646..9422)
/ note="(AB005182) putative DNA binding protein"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3834p"
/ protein_id="AAK87862.1"
/ db_xref="GI:15157248"
/ translation="MHSRYPPMLASMPDFPCLQISLMEWLLAIGNAGNALGRV VEARITLFEQDEPETERKVSFVAIILASAKMAKAGVNDADAVRAFROI FDPDEBAK NVARLYNLARQOVAGYEAVERIAGLCSGNGCMLESLVIGLFIHIAADGLIHRE LAFLGRABIFRITEDHFTETIMARVHMDGRDPYRLVGSPPDDFIDIKRYASLVAE HHPDKLIARGVMEHMAANERMAALNAVAIAEKERRA"
/ complement(9385..10929)
/ gene="AGR_C_3836"
/ complement(9385..10929)
/ note="(AF246209) pPl-phosphofructokinase"
/ codon_start=1
/ trans1_table=11
/ product="AGR_C_3836p"
/ protein_id="AAK87863.1"
/ db_xref="GI:15157249"

```

Alignment Scores:

| Pred. No.: | Length: | Matches: | Mismatches: | Indels: | Gaps: |
|------------------------|---------|---------------|-------------|---------|-------|
| 3,266-86 | 11715 | 246 | 71 | 101 | 7 |
| Score: | 1292.50 | Conservative: | 71 | | |
| Percent Similarity: | 74.59% | Mismatches: | 101 | | |
| Best Local Similarity: | 57.88% | Indels: | 7 | | |
| Query Match: | 56.59% | Gaps: | 2 | | |

```

US-10-701-200-6 (1-437) x AEO08127 (1-11715)
QY 8 HisLeuThrAlaAspIle-----ArgpheCyHisIleTrpPheLeuAenphe 23
    |||||  |||  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9649 CACCTTACACGCGGCTGTCGATCCATTCGCCAGACATGCTTTGGGTATACGAA---- 9705
QY 24 AenpheTyThrLeuMetAenlysProlyslsValAlaIleLeuThrAlaGlyIleu 43
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9706 -----AGAGAGACCATGATGCGAAACAGAAAGTCGCAATGCTGACCGCGGTGCGCTC 9759
QY 44 AAlpTrocysLeuAnsSerAlaIleGlySerLeuIleGlyIleuArgTyThrGluIleAapPro 63
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9760 GGGCCCTGCTCTCTTCGGCGCGCTGATTCGATTCAGCGGCTATACGACATCGCGCCT 9819
QY 64 SerIleGluIleIleCyetyArgGlyIleTyIleGlyLeuLeuLeuGlyAaspSerTy 83
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 9820 GAAATCGACATATCGCTTACCGCTCCGTTATCAGCGGCTGCTTCTGGCGACGCTTATC 9879
QY 84 ProValThrAlaGluValArgIleTyIleValIleuGluIleArgPheGlyIleSerVal 103
    |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

Db 9880 GAGATCACCAAGATATGCGCAAAAGGCGCATCTACGTCATCGTTCAGCGCGCTCGCCG 9939
Qy 104 11eg1yaensearxvallyleuthrhenvallyaspyvallyalarglyleuval 123
Db 9940 ATCGGCAACAGCGCGGTCAAGTCACCAATGCGCGATTGGCCAAAGCGCGCTCGTC 9999
Qy 124 1ygs1ug1ygluaspProglInlyval1alalaaapglInleuvallyaspglyvalaap 143
Db 10000 AAGGAAGCGCAAACTCCGTGGCGCTCGCGCTGGAAGCGCTGCGCATCGCATCAC 10059
Qy 144 11eleuH1sth11g1ygllyaspaapThraentrralalalaaapleuvalalape 163
Db 10060 ATTCTCCATCTACTGCGTGGCCACACACCAACACCGCGCGCATCTTCGCCCTAT 10119
Qy 164 1eua1aargnaamnytyrlyleuthrval11eg1yleuProlysthThralaspaanap 183
Db 10120 CTCGGCGCAACGAGTATGACTTCACTCGCGCTGCGCGCTGGAAGCGGTGATATAGAC 10179
Qy 184 1alPhepco11elyglInserleu1yala1tptthralalag1ug1n1yala1rgtyr 203
Db 10180 GTCGTGCGCATCAACAGTCCGTGCGCGCTGGAAGCGCGCATCGCTCGCTTC 10239
Qy 204 1PheMeaenvalval1alag1uamna1aaenProxargMetleu11e1al1ag1uval 223
Db 10240 TTGCAATGTTCAGCAACGACGACGCGCGCAACCACTTCGTCTATCCACGATGC 10239
Qy 224 1etg1yargnaCyeg1yTtpleuthralalalathralag1n1y1argyleu1eu 243
Db 10300 ATGGCGCGCATGGCGGTGGCTGACCGCGCGCATCGCGCTATATCCAGAAACC 10359
Qy 244 1aspa1galag1utrtpleuProglInleu1yaleuthrarg1u1serTyrg1u1al1a1a 263
Db 10360 CGCGGCAACGACTATGTCAAGCGCTGATGAACACCGCAANTGAAGAACATCAACGCGC 10419
Qy 264 1alPhevalProgl1uMeal1ar1leapleu1u1ag1u1al1y1argyleu1arg1u1al 283
Db 10420 ATCTACCTCGCGGAATGCGCTTCGACATCGAGCGCGGAAGCGCGCTCAAAAGGTC 10479
Qy 284 1etasp1yval1aspCyaval1aen1lepeval1serg1u1y1alag1u1al1a1a1e 303
Db 10480 ATGGAACAAACATGGCTATGTACGCTGTCGTTTCGAAAGTGCGGCGCTCGACCGCATTC 10539
Qy 304 1al1al1ag1u1e1al1al1y1e1g1n1u1al1Pro1arg1asp1al1Pheg1y1h1le1y 323
Db 10540 GTCGCGCGAGCGGAAGCTGCTGCGCAAGCGGTCAAGCGCGCATCGCTCGCATGGAAG 10599
Qy 324 1eua1a1al1a1aspProgl1y1tPheg1y1ug1n1Phe1al1ag1u1e1al1e1y1a1a 343
Db 10600 ATCGACACCATCAATGTTGGCGGTGTTCCAGAAAGCATTCGCGCTCATCGCGCGC 10659
Qy 344 1glu1y1tr1e1u1val1g1n1y1serg1y1y1rPhe1al1arg1al1aser1a1ser1a1a1ap 363
Db 10660 GAGCGCTCATGTCGAGAAAGTCGGGCTATTTCCGCCGCGCGCGCAATGGCGAGC 10719
Qy 364 1apMe1a1rg1e1u11e1y1serCy1a1a1ap1e1u1al1g1u1Cy1al1a1Phe1arg1g1u 383
Db 10720 GATCTTCGCTCATTCAGGCGCATGTCGTCGCGGTGGAAGCGCGCTCATATAGGTC 10779
Qy 384 1serg1y1al11eg1y1H1aspg1u1e1paan1y1aenval1leu1arg1al11eg1u1Phe1ro 403
Db 10780 TCCGGGGTCAACGGTCATACCAAGACCAAGATGGCAATTCGCGACCTTCAGTTTCG 10839
Qy 404 1arg11e1y1g1y1y1y1as1p1ro1Phe1aen11e1a1p1tr1Phe1aen1ser1e1u1ser 423
Db 10840 CATTATCAAGCGCGGAAGCATTTCCAGCTCTCGCAAAATGTTTACCAAGATGAGAA 10899
Qy 424 1glu11e1g1y1g1n1Pro 428
Db 10900 CACGTGCGTCAAGCG 10914

```

```

RESULT 14
AE009161
LOCUS          AE009161          11807 bp          DNA          linear          BCT 28-MAY-2004

```

```

DEFINITION  Agrobacterium tumefaciens str. C58 circular chromosome, section 187
of 256 of the complete sequence.
ACCESSION  AE009161 AE008688
VERSION    AE009161.1 GI:17740565
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kiteajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Boye Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmeri, A., Raymond, C., Rouse, G.,
Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Kreppan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Neeter, E.W.
The genome of the natural genetic engineer Agrobacterium
tumefaciens C58.
Science 294 (5550), 2317-2323 (2001)
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS
2 (bases 1 to 11807)
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kiteajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Boye Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmeri, A., Raymond, C., Rouse, G.,
Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Kreppan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.
and Neeter, E.W.
Direct Submission
Submitted (27-SEP-2001) Department of Microbiology, University of
Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
98195-7242, USA
FEATURES
source
1..11807
Location/Qualifiers
/organism="Agrobacterium tumefaciens str. C58"
/mol_type="genomic DNA"
/strain="C58"
/isolate="U. Washington"
/db_xref="taxon:176299"
384..785
/gene="Atu2104"
384..785
/gene="Atu2104"
795..1391
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/codon_table=11
/product="conserved hypothetical protein"
/protein_id="AAL43095.1"
/db_xref="GI:17740566"
/translation="MTIYAAGAPAFSLYSIGLVCFALFSTIYTRLTPEHEVELIKAG
NISAVAFILGILGIFSLAASAAHSVSLPDVIVMAAIGLVQVLAFLAKLTWMDLH
LKIRGNVAAGLMSGINVIGTLNACMAI"
795..1391
/gene="Atu2105"
795..1391
/gene="Atu2105"
/note="putative; ORF located using Glimmer"
/codon_start=1
/codon_table=11
/product="hypothetical protein"
/protein_id="AAL43096.1"
/db_xref="GI:17740567"
/translation="WHLDVTEHPRLALGTIAIGLFMAFSTEEVLYTSVDKCVASGV
YPKQAASSEDANQHHLTATPRFRNVACEMETYGANCAQLQSVNSDSSSTSAFVP

```

gene
1938
/gene="gap"
/note="synonym: Atu2106"
1939..1938
/gene="gap"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="glutathionylspermidine synthetase (N-terminal)"
/protein_id="AAI43097.1"
/db_xref="GI:17740568"
/translation="MKRIIVPPRPMPREKANA VGS FHEHMGEPYVDDA YCFGLEE IENITIEEPQI LHMCDMDLAD IVNSEAL TFLAI PRDHVIRSMRGDRIL YGRR FLADIYHGAKLL EYNADPTSPVFETGVFYQWVLTDVALGTLP AQDTQGFNSVG EALLI EARGPSRMRMVRIR KSWPAL"
complement(1933..3471)
/gene="tnp"
/note="synonym: Atu2107"
complement(1933..3471)
/gene="trp"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="tsg family transposase"
/protein_id="AAI43098.1"
/db_xref="GI:17740569"
/translation="MYSIADRLAVEL YIRLGKRLNTTIRQLGTPTKNALRGMYREVY OHLDLTQVARAPKYSEARQA LAEHFRTHDRCISGTMRALGPGRGTLTA WRAEAF PEARTSMVRSHWHPGISEERQAGYIGLCSDGESAOYA DGLVS RPTLSMKOOLLG HEASMSKRKKANPKYPEREELEKLELQSRVDLOJEDHLKKA MELKKGGVOD LITSNEKQLIDALKVEYRLPELLAOLRIARSS FYHARMCLADKAAVRSIAEI FRANRCYCVRRLQASIAKRSYSEKVQRKLMOGBHLAVARP RP RRSGTSIGSIA PERLIIRDPAHQANVKMLTDLITERQIPAGKYLS ITDCFDGWISISTQDPAGI CEFVIMFDALGVANGGEERP IHSDRGAHYRWGMLTRISEARLVASMRSKCQSDNA A CEFFPRLATLETFLFPDMKATIIEQIFVAEVDA YIRWNEKR IKISLGSLSPEVROSL GLUL"
3564..4160
/gene="gap"
/note="synonym: Atu2108"
3564..4160
/gene="gap"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="glutathionylspermidine synthetase (C-terminal)"
/protein_id="AAI43099.1"
/db_xref="GI:17740570"
/translation="WYLMDCAIQAGHRAPFLDIRDIG DASGRVADQSRIEFCFLG YPMERHRSFLANUPAESGITIEPWAKIALSNKQLPMLNQPHNHENLLPSFPDD PELVELQDVVPPLSRREGENVSLYRNREILISAGSYGKG YFOAAPAYFKSEFGXY ALVGSWIQDVRSCGIGIREDVSPITANLSRYVPIIEG"
complement(4188..4589)
/gene="Atu2109"
complement(4188..4589)
/gene="Atu2109"
complement(4188..4589)
/product="RNAase P RNA (rnpB)"
/note="Found by BLASTn"
complement(4645..5835)
/gene="sola"
/note="synonym: Atu2110"
complement(4645..5835)
/gene="sola"
/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="sarcosine oxidase"
/protein_id="AAI43100.1"

```

/db_xref="GI:17740571"
/translation="MSGHFNIVYGRGMGMAAARHIAEYVDGVALIGPEPADIKSH
QGVASHYDEAKIRTRIDGDDWALLNRSIARYADIARSQVEFYAPVCLMVGPER
GGNAPFDVDVLRAAARGLGVSTELGDGSLKSRPFYFSEEGCEVFERDNAGYNNPRA
LVKQALIVKAKAGVTLIDDIYVSTRBEDGRASVOTASGAAYTARVLVAAGFSITRD
LILPKPGADYVARTAFVFPFIDDEGLGGYAGMPSLIYEPHDTTKHIIILPPRYPDGKY
YKICGPPDDKRVSDPEIREMFRSGRESVRDLSTVGTIVSYVHSHVSMACVY
SKTSGYPAIGFTASPRIAVLTGNGTAAKSDIEIGRLGALLNDGKIADDAESTDRK
PEFL"
gene
    complement (5895) . 7082)
    /gene="gola"
    /note="synonym: Atu2111"
    complement (5895) . 7082)
    /gene="gola"
    /note="identified by sequence similarity; putative; ORF
    located using Blastx/Glimmer"
    /codon_start=1
    /transl_table=11
    /product="barcosine oxidase"
    /protein_id="AA143101.1"
    /db_xref="GI:17740572"
    /translation="MPDPKYIVVAGMNGMAAAHLSVOTDGVALLIGPAEADRKTHQ
    GVFSSHYDVARITRGFDGDPYMAELARSIRRYEIAKSGIRFTEGACLITENGKG
    LAGVYSRALSSADRILGVETLGAELALRSPFPALPDHGSFEARNNGHINPRALL
    VKACQALAEAGCGRLVRETAHIDTSSRVWMTRECATTTAEKVIYAAAGFTTMALL
    LPSVDMAATGRTIVFPELDEARQALFSAMPSTVLAETEDIVYIILPVRYPDGKYV
    LKIGESBKGRLFTLVQAVDMFSDGTPDEVEFLTKKALSLMPALACPVTSQVSA
    ITCGYPYIGYQSSNIAVLTGNGFNFAKSDIEIGRLGALLNDGKLTDEFALEMSF
    VFV"
gene
    complement (7175) . 7714)
    /gene="gltc"
    /note="synonym: Atu2112"
    complement (7175) . 7714)
    /gene="gltc"
    /note="identified by sequence similarity; putative; ORF
    located using Blastx/Glimmer"
    /codon_start=1
    /transl_table=11
    /product="soluble lytic transglycosylase"
    /protein_id="AA143102.1"
    /db_xref="GI:17740573"
    /translation="MNRPVAVAVAGVVALASQAGISPAKDQDYKTVTLTFVPRKGC
    QIPETKSLPASKVIGSDLIYKTAKRGIWPNLHAIVSVESKFNPARSAGEVGLIM
    YIKETARPMGFGATKATLYDPETNRMQVYATAHQLGGGVCSITLLYNNAGHATN
    RMNPVSRKCGKVALAS"
    complement (7962) . 8741)
    /gene="Atu2113"
    complement (7962) . 8741)
    /gene="Atu2113"
    /note="identified by sequence similarity; putative; ORF
    located using Blastx/Glimmer"
    /codon_start=1
    /transl_table=11

Alignment Scores:
Pred. No.:      3,29e-86      Length:      11807
Score:          1292.50      Matches:      246
Best Local Similarity:  74.59%  Conservative:  71
Query Match:      57.88%  Mismatches:    101
DB:              56.59%  Indels:        7
                  Gaps:          2

US-10-701-200-6 (1-437) x AE009161 (1-11807)

QY      8 HtsleuThrrAlaAapPile-----ArpGheCyshIstTpPhpheLeuAanPhe 23
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      9741 CACCTACACGGCGGCTGTCATCATTTGCCAGACATGCTCTTGGGTATACGCAAA-- 9797

QY      24 AsnPhetyrThrLeuMetAenLylProLylsYerPalAlaIleLeuThrrAlaGlYleu 43
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      9798 -----AGGAAGCCATATATGCCAAGAACGAAAGTCCGAAATGCTGACCGCGGCTGCCTC 98511

44 AlaProCyLeuAenSerAlaIleGlySerLeuIleGluArgTyrThrGluIleAapPro 63

```

```

Db      9852 GCGCCCTCTCTCTCGCCGCGGATGCGCTGATCGAGCCGCTATGAGCACTCCGCT 9911
Qy      64 SerilegilelleCysEtyrArgGlyGlyTyrLysGlyLeuLeuLeuGlyAspSerTyr 83
Db      9912 GAATCGACATATGCTTACCGCTCCGGTTATCGAGGCGTCTCTGGCGAGCGTATC 9971
Qy      84 ProValThrAlaGluValArgLysAlaGlyValLeuGlnArgPheGlyLysSerVal 103
Db      9972 GAGATCAACCAAGATATGCGCAAAAGGCGCATCTACTGATCGATTCAGCGCGCTCGCG 10031
Qy      104 IllegiAaenSerArgValLysLeuThrAsnValLysAspCysValLysArgGlyLeuVal 123
Db      10032 ATCGGCAACAGCGCGGTCAAGCTGACCAATGCCGCGGATTCGCCAAGGCGCGCTCGTC 10091
Qy      124 LysGluGlyLysAspProGlnLysValAlaAlaAspGlnLeuValLysAspGlyValAsp 143
Db      10092 AAGGAAGCGCAATCCGTGCGCGTCCGCTGAGACGCTGCGCGCTGACCGCATCAC 10151
Qy      144 IlleuHsthrIleGlyLysAspAspThrAsnThrAlaAlaAspLeuAlaAspHe 163
Db      10152 ATTCCTCACTATCGTGTGCGACGACCAACACACCGCGCGGACCTTCCGCTCAT 10211
Qy      164 LeuAlaAspAsnThrArgLysLeuThrValIleGlyLeuProLysThrValAspAsnAsp 183
Db      10212 CTCGGCGCCCAAGATATGACTCCGCTGCGCGCTGCGAAGCGGTGATATGAC 10271
Qy      184 ValPheProIleLysGlnSerLeuGlyAlaTPTThrAlaAlaGlnGlnValArgTyr 203
Db      10272 GTCGTCGCGATCAAGACGTCGCTGCGCGCTGACGCGCGCGGAGTCCGCTTCC 10331
Qy      204 PheMetAsnValAlaGlnGlnAsnAlaAsnProArgMetLeuIleValHisGluVal 223
Db      10332 TTCGCAATGTGACGACAGAGCGCGGCGCAAAACCTTGCTCATCCAGCAAGTC 10391
Qy      224 MetGlyArgAsnGlyTyrLeuThrAlaAlaThrAlaGlnGlnTyrArgLysLeuLeu 243
Db      10392 ATGGGCGCGCATGCGGTGCGTGCACGCGCGCGCGCGCTATATCCAGAAACC 10451
Qy      244 AspArgAlaGluTyrLeuProGlnLeuGlyLeuThrArgGlySerTyrGluValHisAla 263
Db      10452 CGCGGCAACGATATGCGAAGCGCTGATGATGAAACCGCAATGAAAGATCGACGCG 10511
Qy      264 ValPheValProGluMetAlaIleAspLeuGlnAlaGlnAlaLysArgLeuArgGluVal 283
Db      10512 ATCTACCTGCGGAAATGCGCTTCGACATCGAGCGGAGCGCGCTCAAGAGGCTC 10571
Qy      284 MetAspLysValAspCysValAsnIlePheValSerGluGlyValGlyValAlaIle 303
Db      10572 ATGGCAAAACATGCTATGTCACGCTGCTGTTCCGAAAGTGCAGGCGCTCGACGCGCATC 10631
Qy      304 ValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGlyHisIleLys 323
Db      10632 GTCGCGGAGCGGGAAGCTGCTGCGCAAGCGCTCAAGCGGAGCGCTTCGCGCATGGAAG 10691
Qy      324 LeuAspAlaValAsnProGlyLysTyrPheGlyGlnGlnPheAlaGlnMetIleGlyAla 343
Db      10692 ATCGACACCATCAATGTTGCGCGCTGTTCCAGAGCAATTCGCGCGCTCATCGCGCGC 10751
Qy      344 GlnLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAlaAsp 363
Db      10752 GAGCGCTCATGCTGAGAAAGTGCAGGCTATTCGCGCGCTCGCGCGCGCAATGCGGAC 10811
Qy      364 AspMetArgLeuIleLysSerCysAlaAspLeuAlaValAlaGlyCysAlaPheArgGlu 383
Db      10812 GATCTTCGCTCATCAAGCGCATGCTGCTGCGCGCTGCAAGCGCGCTCATATAGCTC 10871
Qy      384 SerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIleGluPhePro 403
Db      10872 TCCGGGGTCAACCGGTATGACGAAAGCAGAAATGCGAGCAATTCGCGGCACTTCGCG 10931
Qy      404 ArgIleLysGlyLysProPheAsnIleAspThrAspThrPheAsnSerMetLeuSer 423

```

```

Db      10932 CGTATCAAGGCGGAAAGCATTTGACCTCTCTGCAAAATGTTTACCGAAGTATGAA 10991
Qy      424 GlnIleGlyGlnPro 428
Db      10992 CACGTGCGTCAAGCGC 11006

RESULT 15
SCU51728
LOCUS SCU51728
DEFINITION Streptomyces coelicolor phosphofructokinase ATP-PFK gene, complete cds.
ACCESSION U51728
VERSION U51728.1 GI:1931572
KEYWORDS SOURCE
ORGANISM Streptomyces coelicolor
REFERENCE Streptomyces coelicolor
AUTHORS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
TITLE 1 (bases 1 to 1549)
Alves,A.M., Euverink,G.J., Bibb,M.J. and Dijkhuizen,L.
Identification of ATP-dependent phosphofructokinase as a regulatory step in the glycolytic pathway of the actinomycete Streptomyces coelicolor A3(2)
JOURNAL Appl. Environ. Microbiol. 63 (3), 956-961 (1997)
MEDLINE 97208211
PUBMED 9055413
REFERENCES 2 (bases 1 to 1549)
AUTHORS Da Costa Rodrigues Alves,A.M., Euverink,G.J.W., Bibb,M.J. and Dijkhuizen,L.
TITLE Direct Submission
JOURNAL Submitted (19-MAR-1996) Alexandra M. Da Costa Rodrigues Alves, Microbiology, RUG-BiologicalCenter, Kercklaan 30, Haren, Gr. 9750NN, Netherlands

FEATURES
source 1..1549
/organism="Streptomyces coelicolor"
/mol_type="genomic DNA"
/strain="1109"
/db_xref="taxon:1902"
160..1188
/codon_start=1
/transl_table=11
/product="phosphofructokinase ATP-PFK"
/protein_id="AAC45135.1"
/db_xref="GI:1931573"
/translation="MKVGVLTGGGDCPELNAIVAVKVGQYEDTFGRDWRGP
LEGDTVPIDIPAVRGILPRGGTVLSSSTNPLKQDGRIRIKDNLAALGVEALITIG
EDTIGVATRIADEGVPCVGVKTDNDLSLTDYFPGDTAVGITAIDRLHTAES
HMRVIVYVWRHRGMIMLHSGIAGAAVILIPROPRVEQVCSVTSRPSVAPIV
VVARGAMPRDDMDVKDLSLYGHRVLSGVEMLAKQIERKTGNEATTVLGVORG
GTPSAFDRWLATRFGLAHDVCVHDGDFKMAVLRGTDIVRPVIAETARLKTVDALY
EEVGVFPG"

ORIGIN
Alignment Scores:
Pred. No.: 5,566-14 Length: 1549
Score: 311.00 Matches: 105
Percent Similarity: 43.86% Conservative: 63
Best Local Similarity: 27.42% Mismatches: 137
Query Match: 13.62% Indels: 78
Gaps: 17

US-10-701-200-6 (1-437) x SCU51728 (1-1549)
Qy      34 LysValAlaIleLeuThrAlaGlyLysLeuAlaProCysLeuAsnSerAlaIleGlySer 53
Db      163 AAGGTGGAAGTACTACGAGAGCGGCACTGCCCGGCTCAAGCGCTATCCGGGCGC 222
Qy      54 LeuIleGluArgTyrThrGluLeaAspProSerIleGluIleIleCysEtyrArgGlyGly 73
Db      223 GTCGTCCGCAAGGCGGTCCAG---GAGTACGCGCTACACTTACCGGCTTCCGGAGCGCG 279
Qy      74 TyrLysGlyLeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLysLysAla 93

```

Job time : 6152 secs

```
Db      280 TGGCGGCGCCCTCGAGAGCGCACCGTCCCGCTC-----GACATCCCGCGCTCCGC 333
Qy      94 GlyValIleuGlnArgPheGlyGlySerValIleGlyAsnSerArgValIleLeuThrAsn 113
Db      334 GGCATCTGCCCCGCG---GGCGGACCGCTCTGCGCTCTCCCGCACCAACCGCTCAAG 390
Qy      114 ValIleAspCysValIleArgGlyLeuValIleGlyIleAspProGlnIleValIle 133
Db      391 CAGCGGAGCGCATCCGCGC-----ATC 414
Qy      134 AlaAspGlnLeuValIleAspGlyValAspIleLeuHisThrIleGlyGlyAspThr 153
Db      415 AAGGACACCTCGCCGACCTCGGAGTCCGCTCATCCATCGCGCGGAGACACC 474
Qy      154 AsnThrAlaAlaIleAspLeuAlaIlePheLeuAlaArgAsnAsnTyrglyLeuThrVal 173
Db      475 CTGCGGCTGCGCCACCGCTCCGC-----GACGAGTACCGCGTGCCTGC 519
Qy      174 IleGlyLeuProIleThrValAspAsnAspValPheProIleIleGlnSerLeuGlyAla 193
Db      520 GTCGGCGTCCCAAGACCATCGACCATCTGCGCCACGACCTTACCATCGGTTTC 579
Qy      194 TrpThrAlaIleGlnGlnGlyAlaArgTyrgPheMetAsnValValIleGlnAsnAla 213
Db      580 GACACCGCGCTCGCATCGCACCGAGCCATCGACCGCTGCACACACCGCGAGTCC 639
Qy      214 AsnProArgMetLeuIleValIleGlnValIleMetGlyArgAsnGlyTrpLeuThrAla 233
Db      640 CACATCGCGCTGCTGGTCTGC---GAGGTATGCGCGGACCGCGCTGAGATAGCCCTC 696
Qy      234 AlaThrAlaGlnGlnIleTyrglyIleLeuAspArgAlaGlnTrpLeuProGlnIleGly 253
Db      697 CACTCG-----GCG 705
Qy      254 LeuThrArgGlnSerTyrglyValHisAlaValPheValProGlnMetAlaIleAspLeu 273
Db      706 CTGCGCGCGCGCGCC-----AACGTATCTCATCCCGACGACGCGCTTCGACGTC 756
Qy      274 GlnAla-----GlnIleIleValIleArgIleValIleMetAspIleValIleAspCys 289
Db      757 GAGCAGGTGCTCTCGGAGTACCTCCCGGCTTCGCGGCTCTTACGCGCCGATC----- 810
Qy      290 ValAsnIlePheValSerGlnGlyAlaGlyValIleAlaIleValIleGlnMetGlnAla 309
Db      811 -----GTGCTGCTCGCGCGCGC-----ATGCGCGC 840
Qy      310 LysGlyGlnGlnValProArg-----AspAlaPheGlyHisIleIleValIleAsp 325
Db      841 GACGCGGACATGCTGCTCAAGACGAGTCCGCTGACTCTTACGCGGACGTCGCTGTC 900
Qy      326 AlaValAsnProGlyIleTrpPheGlyGlnGlnPheAlaGlnMetIleGlyAlaGlu 344
Db      901 GGGGTC-----GGCGAATGCTGCGCCAGACATCGAGAAAGCGCACCGCAAGAGCC 954
Qy      345 LysThrLeuValGlnIleSerGlyTyrgPheAlaArgAlaSerAlaSerAsnValAspAsp 364
Db      955 CGACACCACTGC-----CTCGGACCATGTCACCGCGCGGACGCGCCGCTTCGAC 1008
Qy      365 MetArgLeuIleIleAspCysAlaAspLeuAlaValIleGlnCysAlaPheArgArgIleSer 384
Db      1009 CGCTGCTGCGCCACCGCTTCGAGCTGCAAGCGCTGCACTGCGCTG----- 1053
Qy      385 GlyValIleGlyHisAspGlnAspAsnGlyAsnVal-----LeuArgAlaIleGlnPhe 402
Db      1054 -----CAGCAGCGGACTTCGCGAAGATGTCCTGCGCGGACGACGACATC 1101
Qy      403 ProArgIle 405
Db      1102 GTCGCGCTC 1110
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 01:38:40 / Search time 721 Seconds
(without alignments)
3587.971 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVTWPYHUTADIRCFHWF.....FNSMSEIQPKGKRVESH 437

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4390206 segs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODBI=framer+ p2n model -DRV=hlh
-Q=/cgn2_1/USPTO.spool/US10701200/runat.25022005.102832.10300/app.query.fasta_1.583
-DB=N.Geneseq.16Dec04 -QFMT=fastap -SUFIT=frng -MINMATCH=0.1 -LDOFCL=0
-LOOPEXT=0 -UNITS=bits -STRAT=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10701200 @CGN 1.1 470 @runat.25022005.102832.10300 -NCPU=6 -ICPU=3
-NO MAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FEAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.16Dec04:*
1: geneseqn1800s:*
2: geneseqn1800s:*
3: geneseqn2000s:*
4: geneseqn2000s:*
5: geneseqn2000s:*
6: geneseqn2000s:*
7: geneseqn2000s:*
8: geneseqn2000s:*
9: geneseqn2000s:*
10: geneseqn2000s:*
11: geneseqn2000s:*
12: geneseqn2000s:*
13: geneseqn2000s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 2284 | 100.0 | 1311 | 6 | ABL51513 |
| 2 | 2284 | 100.0 | 1311 | 6 | AAD35497 |
| 3 | 2284 | 100.0 | 1311 | 6 | ABK83232 |
| 4 | 1444.5 | 63.2 | 8530 | 4 | AAS59528 |
| 5 | 1444.5 | 63.2 | 8530 | 8 | ACF64457 |

| | | | | | | |
|----|-------|------|-------|----|----------|--------------------|
| 6 | 659 | 28.9 | 552 | 13 | ADS62361 | Ad62361 Bacterial |
| 7 | 306.5 | 13.4 | 1050 | 13 | ADS56121 | Ad56121 Bacterial |
| 8 | 281.5 | 12.3 | 1458 | 3 | ACA46084 | ACA46084 Arabidops |
| 9 | 270 | 11.8 | 1347 | 8 | ACA23197 | ACA23197 Prokaryot |
| 10 | 270 | 11.8 | 1347 | 8 | ACA23197 | ACA23197 Prokaryot |
| 11 | 270 | 11.8 | 1347 | 8 | ACA23197 | ACA23197 Prokaryot |
| 12 | 264.5 | 11.6 | 1624 | 2 | AAQ85984 | AAQ85984 Oryza sat |
| 13 | 263.5 | 11.5 | 2256 | 6 | ADA70655 | ADA70655 Rice gene |
| 14 | 263 | 11.5 | 2256 | 6 | ADA70655 | ADA70655 Rice gene |
| 15 | 262 | 11.5 | 960 | 13 | ADS48162 | ADS48162 Bacterial |
| 16 | 262 | 11.5 | 1978 | 2 | AAQ85982 | AAQ85982 Solanum t |
| 17 | 261.5 | 11.4 | 1490 | 8 | ADA70675 | ADA70675 Rice gene |
| 18 | 261.5 | 11.4 | 1491 | 10 | ADC07917 | ADC07917 Rice DNA |
| 19 | 261 | 11.4 | 1035 | 13 | ADT44510 | ADT44510 Bacterial |
| 20 | 258.5 | 11.3 | 1769 | 3 | ACA44555 | ACA44555 Zea mays |
| 21 | 257.5 | 11.3 | 1503 | 8 | ADA68295 | ADA68295 Arabidops |
| 22 | 257 | 11.3 | 1503 | 8 | ADA68295 | ADA68295 Arabidops |
| 23 | 254.5 | 11.1 | 1550 | 8 | ADA70520 | ADA70520 Rice gene |
| 24 | 251 | 11.0 | 1810 | 3 | ACA44927 | ACA44927 Arabidops |
| 25 | 249.5 | 10.9 | 956 | 13 | ADS57682 | ADS57682 Bacterial |
| 26 | 248 | 10.9 | 1073 | 13 | ADS57682 | ADS57682 Bacterial |
| 27 | 246.5 | 10.8 | 2964 | 13 | ADT47247 | ADT47247 Bacterial |
| 28 | 240.5 | 10.5 | 960 | 8 | ACA48934 | ACA48934 Prokaryot |
| 29 | 240.5 | 10.5 | 963 | 4 | AAS56050 | AAS56050 Salmonell |
| 30 | 240.5 | 10.5 | 963 | 8 | ACA50979 | ACA50979 Prokaryot |
| 31 | 240.5 | 10.5 | 963 | 8 | ACA51971 | ACA51971 Prokaryot |
| 32 | 240 | 10.5 | 1605 | 3 | ACA42127 | ACA42127 Arabidops |
| 33 | 238.5 | 10.4 | 963 | 4 | AAS52688 | AAS52688 E. coli D |
| 34 | 238.5 | 10.4 | 963 | 13 | ACA32277 | ACA32277 Prokaryot |
| 35 | 238.5 | 10.4 | 963 | 13 | ADT48887 | ADT48887 Bacterial |
| 36 | 234.5 | 10.3 | 984 | 8 | ACA54036 | ACA54036 Prokaryot |
| 37 | 234 | 10.2 | 1530 | 3 | ACA48983 | ACA48983 Arabidops |
| 38 | 232 | 10.2 | 1386 | 8 | ACA52344 | ACA52344 Prokaryot |
| 39 | 232 | 10.2 | 37337 | 4 | AAS59518 | AAS59518 Prokaryot |
| 40 | 232 | 10.2 | 37337 | 4 | ACF64447 | ACF64447 Prokaryot |
| 41 | 230.5 | 10.1 | 1164 | 12 | ADJ40285 | ADJ40285 Plant cDN |
| 42 | 229.5 | 10.0 | 885 | 13 | ADS56868 | ADS56868 Bacterial |
| 43 | 229.5 | 10.0 | 960 | 8 | ACA32326 | ACA32326 Prokaryot |
| 44 | 228 | 10.0 | 7972 | 2 | AAV74623 | AAV74623 Staphyloc |
| 45 | 227.5 | 10.0 | 1349 | 13 | ADS56889 | ADS56889 Bacterial |

ALIGNMENTS

| | | |
|----------|--|---|
| RESULT 1 | ABL51513 | standard; DNA; 1311 BP. |
| XX | ABL51513; | |
| AC | ABL51513; | |
| XX | 01-JUL-2002 (first entry) | |
| DT | 01-JUL-2002 (first entry) | |
| XX | Methylomonas pyrophosphate dependent phosphofructokinase DNA SEQ ID:15. | |
| XX | Methylomonas; methanotropic; carbon metabolism; carbon flux pathway; | |
| KW | transaldolase; fructose biphosphate aldolase; KMG/KDPG aldolase; food; | |
| KW | pyrophosphate dependent phosphofructokinase; phosphoglucosylase; feed; | |
| KW | glucose 6 phosphate isomerase; 6-phosphoglucosylase; dehydratase; methane; | |
| KW | glucose 6 phosphate 1 dehydrogenase; enzyme; methanotroph; methanol; | |
| KW | single carbon substrate; single cell protein; polysaccharide; thickener; | |
| KW | isoprenoid; carotenoid pigment; gene; ds. | |
| XX | Methylomonas sp. | |
| OS | Methylomonas sp. | |
| XX | | |
| PH | | |
| FT | Key | Location/Qualifiers |
| FT | CDS | 1..636 |
| FT | | /tag= a |
| FT | | /partial |
| FT | | /product= "pyrophosphate dependent phosphofructokinase" |
| XX | | /note= "no start or stop codons given" |
| XX | | |
| PN | MO200220796-A2. | |

PD 14-MAR-2002.
 XX 28-AUG-2001; 2001WO-US026730.
 PF 01-SEP-2000; 2000US-0229906P.
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.
 PA Koffas M, Odum JM, Norton KC, Ye RW,
 PI WPI; 2002-362250/39.
 DR P-PSDB; ABB09168.
 XX
 PT New polynucleotide encoding a Methylenomonas sp. carbon flux enzyme useful
 PT for altering carbon flow through methanotrophic bacteria, utilized for
 PT production of single cell protein and commercially valuable
 PT polysaccharides.
 PS Claim 2; Page 66; 73pp; English.
 XX
 XX The present invention describes isolated polynucleotides (I) encoding a
 CC Methylenomonas sp. carbon flux enzyme, consisting of: transaldolase;
 CC fructose biphosphate aldolase; KDG/KDPG aldolase; phosphoglucosylase;
 CC pyrophosphate dependent phosphofructokinase; glucose 6 phosphate
 CC isomerase; 6-phosphogluconate dehydratase; and glucose 6 phosphate 1
 CC dehydrogenase enzymes. (i) overexpression is useful for altering carbon
 CC flow through a methanotrophic bacteria, where the bacteria grows on a C1
 CC carbon substrate of methane and methanol, and comprises a functional
 CC Embden-Meyerhof carbon pathway comprising a gene encoding a pyrophosphate
 CC dependent phosphofructokinase enzyme, where the bacteria is preferably
 CC Methylenomonas 16a American Type Culture Collection (ATCC) PTA 2402, and
 CC where the carbon flux gene is: (i) over-expressed on a multicopy plasmid;
 CC (ii) operably linked to an inducible or regulated promoter; (iii)
 CC expressed in an antisense orientation; or (iv) disrupted by insertion of
 CC foreign DNA into the coding region. The manipulated methanotrophs are
 CC useful for the energy efficient conversion of single carbon substrates
 CC such as methane and methanol to commercially useful products in the food,
 CC feed and materials industries, and preferably for the production of
 CC single cell protein, and for producing polysaccharides, used as
 CC chickeners in food and non-food industries, and isoprenoid compounds and
 CC carotenoid pigments of various carbon lengths. The present sequence
 CC encodes a Methylenomonas pyrophosphate dependent phosphofructokinase from
 CC the present invention
 XX
 SQ Sequence 1311 BP; 327 A; 320 C; 366 G; 298 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,78e-229 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conserved: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-701-200-6 (1-437) x ABL51513 (1-1311)
 QY 1 AapvAlvalThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhePhe 20
 DB 1 GATGGTGCATCATGCGCCATCATCACTTAACGGCTGATTCGATTTGTCATGTTT 60
 QY 21 LeuAnpPheAsnPheTyrThrLeuMetAnlyrProLysValAlaIleLeuThrAla 40
 DB 61 CTTAACCTTAACTTCTTACACGCTCATGAACCAACCTAAAGTTGCATACTGACAGCA 120
 QY 41 GlyGlyLeuAlaProCysLeuAnSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
 DB 121 GCGGCGCTTGCGCCCTTGTGTAATTCGCAATCGGTAGTTGATGCAACGTATATCCGAA 180
 QY 61 IleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
 DB 181 ATCGATCTCAGCATGAATCATATTGCTATCGCGGGGCTTAAAGAGCGCTGTGCTGGGC 240
 QY 81 AapSerTyrProValThrAlaGluValArgLysValAlaGlyValLeuGlnArgPheGly 100

DB 241 GATTCCTATCCAGTACCGCCGAAAGTCGTAAAAAGCGGGGTTCGTGCAACGTTTGGC 300
 QY 101 GlySerValIleGlyAsnSerArgValIlybLeuThrAsnValIlyAspCysValIlybArg 120
 DB 301 GGTTCGTGATCGGCAACAGCGCGGTCAAAATTGACCAATGTCAAAAGATCGGTGAAAGC 360
 QY 121 GlyLeuValIlybGlyGlyGluAspProGlnLysValAlaIleAspGlnLeuValIlybAsp 140
 DB 361 GGTTCGTCAAAAGAGCGTGAAGATCCGCAAAAGTCGGCGGTGATCAATTGGTTTAAAGAT 420
 QY 141 GlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaIleAspLeu 160
 DB 421 GGTTCGTATTTCTGCAACCATCGCGCGGTATGATACCATACGACGACGCGATTTG 480
 QY 161 AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrVal 180
 DB 481 GCAGCATTCCTGGCCGCAAAATTAATTAGGACGTGACCGTCAATTGTTACTAAACCGTC 540
 QY 181 AapAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGly 200
 DB 541 GATTAACGACGTATTTCCGATCAACCAATCACTAGGTGCTTGGACTGCGCGCAAGAGC 600
 QY 201 AlaArgTyrPheMetAsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
 DB 601 GCGCTTATTTCAATGAAAGTGTGCGCAAAACAAAGCCAAACCGCATGCTGATCGTA 660
 QY 221 HisGluValMetGlyArgAsnCysGlyTyrLeuThrAlaAlaIleThrAlaGlnLysTyrArg 240
 DB 661 CACGAAGAGATGGCGCGTAACTGCGCGGTGACCGGTGCAACCGCGCAAGCAAGAAATGCGC 720
 QY 241 LysLeuLeuAspAspArgAlaGluTrpLeuProGlnLysGlyLeuThrArgGluSerTyrGlu 260
 DB 721 AAATTTACTGACCGCGCGCAAGTGTGCGCAATGGTGTGACTCGTGAATCTTATGAA 780
 QY 261 ValHisAlaValAlaPheValProGluMetAlaIleAspLeuGluAlaIlybArgLeu 280
 DB 781 GTGCACGGGGTATTCGTCGCAAAATGGCGATCGATCGAAAGCCGAAAGCCAGGCGCTG 840
 QY 281 ArgGluValMetAspLysValIlybAspCysValIlybAsnIlePheValSerGluGlyValAlaGly 300
 DB 841 CCGCAAGAGATGGAACAAAGTGGATTCGCTCAACATCTTCTTCCGAAGTCCGCGCGTC 900
 QY 301 GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 DB 901 GAAGCTATCGTCGCGGAAATGACAGCCAAAGCCAGAAAGTCCCGCGGATCGTTCGCGC 960
 QY 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
 DB 961 CACATCAAACTGAGTGGCGGTCAACCTGTGTAAATGTTGGGGAGCAATTCGCGCAATG 1020
 QY 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 DB 1021 ATAGCGCGGGAAGAAACCTCGTGTCAAAATCGGGAATCTTCCGCGGTCTTCGCTTCC 1080
 QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
 DB 1081 AACGTTGACGACATGCGTTTATCAAAATCGTCCGCGCTTGGCGGTGAGAGCGCGCTTC 1140
 QY 381 ArgArgGlySerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
 DB 1141 CGCCGCGAGTCTGGGTATCGGTCAACAGCAAGCAACCGCAACGTGTGGTGCATTC 1200
 QY 401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspThrAspTrpPheAsnSer 420
 DB 1201 GAGTTTCGCGCATCAAGCGCGCAACCGTTCATATCGACACGCACTGTTCAATAGC 1260
 QY 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
 DB 1261 ATGTTGACGAAATCGGCAAGCTTAAAGCGGTAAAGTCGAAAGTCAGCCAC 1311
 RESULT 2
 AAD35497

ID AAD35497 standard; DNA, 1311 BP.
 AC AAD35497;
 XX
 XX
 DT 07-AUG-2003 (revised)
 DT 25-JUL-2002 (first entry)
 XX
 XX Methylenomonas 16a sp. phosphofructokinase pyrophosphate dependent gene.
 DE
 XX Carotenoid; isopentenyl pyrophosphate; antheraxanthin; asaxanthin; diet;
 KM anti-oxidant; steroid; flavour; fragrance; electro-optic application;
 KM aquaculture; enzyme; phosphofructokinase pyrophosphate; gene; ds.
 XX
 OS Methylenomonas sp.
 FH Key Location/Qualifiers
 FT CDS 1..1311
 FT /tag a
 FT /product= "Phosphofructokinase pyrophosphate"
 FT /note= "CDS does not include start and stop codon"
 FT /partial
 XX
 PN MO200218617-A2.
 XX
 XX 07-MAR-2002.
 XX
 XX 04-SEP-2001; 2001WO-US027420.
 XX
 XX 01-SEP-2000; 2000US-0229858P.
 XX 01-SEP-2000; 2000US-0229907P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 PI Brzostowicz PC, Cheng Q, DiCosimo DJ, Koffas M, Miller ES;
 PI Odom JM, Plicataggio SK, Rouviere PB;
 XX
 DR WPI; 2002-351711/38.
 DR P-PSDB; AAE22299.
 XX
 XX Producing carotenoid compounds e.g. antheraxanthin and asaxanthin, by
 PT using microorganisms having a nucleic acid molecule encoding enzymes in
 PT the carotenoid biosynthetic pathway and which metabolize single carbon
 PT substrates.
 XX
 XX Example 2; Page 103-104; 156pp; English.
 PS
 XX The invention relates to a method for producing carotenoid compounds. The
 CC method comprises a transformed metabolizing host cell, comprising
 CC suitable levels of isopentenyl pyrophosphate and a nucleic acid molecule
 CC encoding an enzyme in the carotenoid biosynthetic pathway, under the
 CC control of regulatory sequences, and contacting the host cell with carbon
 CC substrate to produce a carotenoid compound. The method is useful for
 CC producing carotenoid compounds such as antheraxanthin and asaxanthin, by
 CC using microorganism having a nucleic acid molecule encoding enzymes in
 CC the carotenoid biosynthetic pathway and which metabolize single carbon
 CC substrates. The carotenoids have potent anti-oxidant properties useful in
 CC diet, and aquaculture elements. The carotenoids are also useful as
 CC intermediates in the syntheses of steroid flavours and fragrances and
 CC compounds for potential electro-optic applications. The present sequence
 CC is Methylenomonas 16a sp. phosphofructokinase pyrophosphate dependent
 CC enzyme gene used in the invention. (Updated on 07-AUG-2003 to correct OS
 CC field.)
 CC
 SO Sequence 1311 BP; 327 A; 320 C; 366 G; 298 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,78e-229 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-701-200-6 (1-437) x AAD35497 (1-1311)
 QY 1 AspValValThrTrpProTyrHisLeuThrAlaAsp11eArpGheCyHisTrpPhePhe 20
 DB 1 GATGTGTCACATGAGCCCTATCATCTTAACGGCTGATTCGATTTGCTATTCGTTT 60
 QY 21 LeuAsnPheAsnGheTyrThrLeuMetAsnLysProLysLysValAla11eLeuThrAla 40
 DB 61 CTTAACTTAACCTTACACAGCTCATGAACTAAACCTAAAGTTCCAACTACGACAGCA 120
 QY 41 GlyGlyLeuAlaProCyLeuAsnSerAla11eGlySerLeu11eGlyuArgTyrThrGlu 60
 DB 121 GGGGCTTGCGCTTGTGTTGAATTCGCAATCGTAGTTGATTCGACAGCTTATCCGAA 180
 QY 61 11eAspProSer11eGlyuLeu11eGlyTyrArgGlyGlyTyrArgGlyLeuLeuGly 80
 DB 181 ATCGATCCCTAGCATTAATAATCATTTCTATCCGGGCGGTTATTAAGCCCTGTTCGCGC 240
 QY 81 AspSerTyrProValThrAlaGluValArgLysLysAlaGlyValLeuGlnArgPheGly 100
 DB 241 GATTCTTATCCAGTAACGCCCGAAGTCGTAAAAAGCGGGTGTCTGCAACGTTTGGC 300
 QY 101 GlySerVal11eGlyAsnSerArgValLysLeuThrAsnValLysAspCyValLysArg 120
 DB 301 GGTTCGTGATCGCAACAGCCGCTCAATATGACCAATGTCAAGACTGCGTGAACGC 360
 QY 121 GlyLeuValLysGlyGlyGlyuAspProGlnLysValAla11eAspGlnLeuValLysAsp 140
 DB 361 GGTTCGTGATCGCAACAGCCGCTCAATATGACCAATGTCAAGACTGCGTGAACGC 420
 QY 141 GlyValAsp11eLeuHisThr11eGlyGlyAspAspThrAsnThrAla11eAspLeu 160
 DB 421 GGTGTGATATTCGCAACCATCGCGGATGATACCAATACGCGACAGCGGATTTG 480
 QY 161 AlaAlaPheLeuAlaArgAsnSerTyrGlyLeuThrVal11eGlyLeuProLysThrVal 180
 DB 481 GCAGCATTCCTGCGCCGAAATATTAATGACATGACCTGCTTACTTAACCTAAACCGTC 540
 QY 181 AspAsnAspValPheProLysLeuGlnSerLeuGlyAla11eThrAla11eGlnGly 200
 DB 541 GATTAACGAGTATTTCCGATCAGCAATCACTAGTGCTTGGATCCCGCGAGCAAGGC 600
 QY 201 AlaArgTyrPheMetAsnValAla11eGlyAsnAsnAlaAsnProArgMetLeu11eVal 220
 DB 601 GCGCGTATTTATGTAACGCGGTGGCGGCAAAACCAACCCACGACATGCTGATCGTA 660
 QY 221 HisGlyValMetGlyArgAsnCyGlyTyrPheuThrAla11eThrAla11eGlnGlyTyrArg 240
 DB 661 CACGAAGTATGAGCGCGTAATCTGCGCTGCTGACCGCTGCAACCGCGAGGAATTCGC 720
 QY 241 LysLeuLeuAspArgAlaGluTyrLeuProGlnLeuGlyLeuThrArgLysSerTyrGly 260
 DB 721 AAATTACTGACCGTCCGAGTGTGTCGGAATGGGTTTGCTGTAATTTATGAA 780
 QY 261 ValHisAlaValPheValProGlnMetAla11eAspLeuGlnAlaGluAlaLysArgLeu 280
 DB 781 GTGCAAGCGGTATTTGCTTCGGAATGGGATGACCTGGAAGCGAAGCGAGCGCTG 840
 QY 281 ArgGlyValMetAspLysValAspCyValAsn11ePheValSerGlyValAlaGlyVal 300
 DB 841 CGGGAAGTATGACCAAGTCGATTCGTCACATCTTCGTTCCCAAGGTGCGCGCTC 900
 QY 301 GluAla11eValAlaGlnMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
 DB 901 GAAGCTATCGTCGCGAATATGACGCAAGCAAGGCGCAAGTGCCTGCTTCGCGC 960
 QY 321 HisLeuLeuAspAlaValAsnProGlyTyrTrpPheGlyGlyGlnPheAlaGlnMet 340
 DB 961 CACATCAAACTGATCGGTGCAACCTGTGAATATGTTGGGAGCAATTCGCGAGATG 1020
 QY 341 11eGlyAlaGlnLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
 DB 1021 ATAGGCGCGGAAAAAACCCCTGTGTAACAAAATCGGAGTACTTCGCGCTGTCTGCTTCC 1080

| | | | |
|----------|--|------|--|
| Oy | | 361 | AenValAAPAPeMWArXgenUlelySeScYgAlAAsPLeMaLAValGlUCYSALApPe 380 |
| Dd | | 1081 | AACGTTCAGCACAATCGCTTGGATTCAAAATGTCGCCACTTGCCGTGCAGTGGCGTTTC 114 |
| Oy | | 381 | ArgAtGSgluSerglVallIegYHtASrPGluASPAnSGlvANvalIeuATrgAIJile 400 |
| Dd | | 1141 | CCCCCGAGNCTcTGccctGAgttcgcTAcCaAGAACAacGaCcgaactGttcgTcggatTC 120 |
| Oy | | 401 | GluPhEPProARgIlElYvGIylYSPProPhEAEnllEasPTraSpTrpheaNSer 420 |
| Dd | | 1201 | GAGTTCCTCGCATCATGAGGCGCGCAAACCCTTCAATATGCACACcaCTGTTgTCCAATAgC 126 |
| Oy | | 421 | MettAusSerGUilIEGLYGlnPrOlySGlYglYLVSVALGVAlSErtHS 437 |
| Dd | | 1261 | ATGTTGAGCGAATAcGGCCAGcCTTAAGGCGGTAAAGTCAAGTCAGCCAC 1311 |
| RESULT 3 | | | |
| ID | ABK83232 | | |
| XX | ABK83232 standard; DNA; 1311 BP. | | |
| AC | ABK83232; | | |
| DT | 27-AUG-2002 (first entry) | | |
| DE | High growth methanotrophic bacterial strain DNA #3. | | |
| KW | High growth methanotrophic bacterial strain; C1 carbon substrate; gene; | | |
| KM | methane; methanol; Embden-Meyerhof carbon flux pathway; 16S rRNA; ds; | | |
| KW | pyrophosphate dependent phosphofructokinase; nitrogen-containing compound; | | |
| KW | ammonia; nitrate; nitrates; nitrogen; pigment; oxygen; landfill; | | |
| KW | methane-containing environment; waste water treatment system; isoprenoid; | | |
| KW | nitrous oxide; terpenoid; animal feed; carotenoid; exopolysaccharide. | | |
| OS | Methylomonas sp. | | |
| PN | WO200220728-A2. | | |
| PD | 14-WAR-2002. | | |
| PF | 28-AUG-2001; 2001MO-US026827. | | |
| PR | 01-SEP-2000; 2000US-022985BP. | | |
| PA | (DUPO) DU PONT DE NEMOURS & CO E I. | | |
| PI | Koffas M, Odom JM, Schenzle A; | | |
| DR | WP1; 2002-454358/48. | | |
| XX | P-PsDB; ABG61553. | | |
| PT | New high growth methanotrophic bacterial strain, useful for producing | | |
| PT | single cell proteins, grows on a C1 carbon substrate, and comprises a | | |
| PT | functional gene encoding in Embden-Meyernof carbon pathway. | | |
| PS | Claim 29; Page 85; 157pp; English. | | |
| CC | The invention relates to a high growth methanotrophic bacterial strain, | | |
| CC | which grows on a C1 carbon substrate e.g. methane and methanol, and | | |
| CC | comprises a functional Embden-Meyernof carbon flux pathway comprising a | | |
| CC | gene coding a pyrophosphate dependent phosphofructokinase enzyme or a 16S | | |
| CC | rRNA. The bacterial strain is useful for the production of single cell | | |
| CC | protein and for the biotransformation of a nitrogen-containing compound, | | |
| CC | e.g. ammonia, nitrate, nitrite or nitrogen. It is also useful for the | | |
| CC | production of a feed product comprising a protein, carbohydrates and a | | |
| CC | pigment and for reducing oxygen demand, for removing nitrates and | | |
| CC | nitriles in methane-containing environments such as landfills, waste | | |
| CC | water treatment systems or anywhere that methane, oxygen and nitrates are | | |
| CC | present. The bacterial strain of the invention can be used as a | | |
| CC | identifying agent for the conversion of nitrate or nitrite to nitrous | | |
| CC | oxide with methane or methanol as a carbon source. It is also used in the | | |
| CC | production of biomass including proteins, carbohydrates and a wide | | |
| CC | variety of pigments (particularly for isoprenoid pigments) for the purpose | | |

[illegible]

QY 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluGluVal 300
DB 841 CGGAAAGTATGATGACAAAGTCGATTCGTCACATCTTGGTTCCGAAGGTGCGCGTC
QY 301 GUAAlIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
DB 901 GAAGCTATGCTCCCGAATAATGACGGCCAAAGCCAGGAAGTCCCGCGATGCTTCGCGC
QY 321 HsIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGlnMet 340
DB 961 CACATCAAACTGATCGGTCGATCGATCCCTGGTAATGTTGCGCGACGCAATTCGCGCAATG 1020
QY 341 ILeGlyAlaGluLysTrpLeuValGlnLysSerGlyTyrosPheAlaArgAlaSerAlaSer 360
DB 1021 ATAGGGCGGAAAAAACCCCTGATCAAAAATCGGATACCTTCCCGCTGCTTCTGCTTC 1080
QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
DB 1081 AACGTTGACGACATCGCTTGTATCAATCGTCGCGCACTTGGCGGTGCGAGTCCGCTTC 1140
QY 381 ArgArgGluSerGlyValIleGlyHsAspGluAspAsnGlyAsnValLeuArgAlaIle 400
DB 1141 CGCGCGAGCTGCGCTGATCGGTCAAGACGAAACGCGCAACGTTGCGTGCATC 1200
QY 401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspTrpPheAsnSer 420
DB 1201 GAGTTCCCGGATCAAGGCGGCAACGCTTCAATATGACACCGACTGTTCAATAGC 1260
QY 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
DB 1261 ATGTTGAGCGAATCGCGCCAGCTAAAGCGGTAAAGTCGAAGTCAGCCAC 1311

RESULT 4

AAS59528
ID AAS59528 standard; DNA; 8530 BP.

AC AAS59528;

DT 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein encoding DNA #23.

KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant; de.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-019047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (COR1-) CORIXA CORP.

XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'mabonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX Claim 1; SEQ ID NO 23; 10699P; English.

CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC associated DNA sequences are used in the treatment, prevention and
CC diagnosis of medical conditions caused by P. acnes. The disorders include
CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of P. acnes in a
CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for P. acnes proteins. These antibodies can be
CC used to downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU4543-AAU45926 and AAU67500. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

CC Sequence 8530 BP; 1921 A; 2633 C; 2423 G; 1551 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 2,72e-140 Length: 8530
Score: 1444.50 Matches: 274
Percent Similarity: 80.59% Conservative: 54
Best Local Similarity: 67.32% Mismatches: 74
Query Match: 63.24% Indels: 5
DB: 4 Gaps: 1

US-10-701-200-6 (1-437) x AAS59528 (1-8530)

QY 33 LysLysValAlaIleLeuThrAlaGlyLysLeuAlaProCysLeuAsnSerAlaIleGly 52
DB 3962 AAGAAAGTGCACTCCTCAACCGGGCGGATTTGCGCTGCTATCCACCGGATCAGT 4021
QY 53 SerLeuIleGluArgTrpThrGluIleAspProSerIleGluIleLysTrpArgGly 72
DB 4022 GANTTATCAGGCGCTACACCGAGTACGACCTGAGTGCAGATCATCCGCTCAAGCAC 4081
QY 73 GlyTrpLysGlyLeuLeuLeuGlyAspSerTrpProValThrAlaGluValArgLys 92
DB 4082 GGTTACAGAGGCGCTGTCGAAGGAGCTTCCTCGAGTCAACGACCGTTGCCAAGAAT 4141
QY 93 AlaGlyValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThr 112
DB 4142 GCCGAGATCCTCAAGCGTTTGGTGGATCCCGATGGCAACTCGCGGATCAAGCTGACC 4201
QY 113 AsnValLysAspCysValLysArgGlyLeuValLysGluGlyLysAspProGlnLysVal 132
DB 4202 AACGCCCGGACCTCGTAAGCGTGTGTCGTCGCGAAGGCGACGACCGCTCAAGTT 4261
QY 133 AlaAlaAspGlnLeuValLysAspGlyValAspIleLeuHsIleThrIleGlyLysAsp 152
DB 4262 GCGCGGATCGTCTGTTCCGACGCGGTTGACATTTTGCACACCATCGGTGTGACGAC 4321
QY 153 ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTrpGlyLeuThr 172
DB 4322 ACTAATACCAACCGCCCGCATCTGCGCGCTCTCTGAGTAATTAACGCGCTCACCC 4381
QY 173 ValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSerLeuGly 192
DB 4382 GTGTTGGCTCGGAAGACCATCGATACGACGTGTCGCCATTCGTCAATGCTAAGT 4441
QY 193 AlaTrpThrAlaAlaGluGlnGlyAlaArgTrpPheMetAsnValAlaGluAsnAsn 212
DB 4442 GCTTGACGCGCGGAAACAAGTTCCTGCTTGTCTCAAGACATCGTTGCGGACGACAC 4501
QY 213 AlaAsnProArgMetLeuIleValHsGluValMetGlyArgAsnGlyGlyTrpLeuThr 232
DB 4502 TCCGATCAACGATGCTCATGTCACGAAGTCATGGCGGTAACTCGGCTGGCTCACCC 4561

| | | | |
|----|------|--|------|
| QY | 233 | AlAlaIaThraLaGInGluTyrArgGlyLeuLeuAspAgaIaGluTyrLeuProGluLeu | 252 |
| Db | 4662 | GCCTCTACCGCGCCCAATATCCGCGAAAGGCTCATACTCAACAAGTGGCTTACTGAGATC | 4622 |
| QY | 253 | GlyLeuThraArgJuseTyrGluValHlaIaValPheValProGluMetAlaIleAsp | 272 |
| Db | 4622 | GGCGCTGTCCAAAGAGGCTCGGAGATGCCACCCGATATGTCTCCGAGGCTCACATTCGAC | 4681 |
| QY | 273 | LeuGluIaLaGluAlaValArgLeuAspGluValMetAspLysValAspCysValAsnIle | 292 |
| Db | 4682 | CTTAAGGCCGAAGCGCGCAGGCTCTCAACAGGTCATGACGAAAGTTGGCAAGCTCATATC | 4741 |
| QY | 293 | PheValSergJugIaGlyAlaGlyValaGluIaIleValaIaGluMetGlnAlaIleSgIyGln | 312 |
| Db | 4742 | TTCTCTGTCCAGGGGGGCTGGCTGGAGCCATTATTGAGGATGGAAGAGAGCGTCAG | 4801 |
| QY | 313 | GluValProArgAspAlaPheGlyHlaIleLysLeuAspAlaValAsnProGlyLysTrp | 332 |
| Db | 4802 | GAACTCTCCGCGCGATCCCTTTGTGCATCTCAAGCTCGCAAGGTGAAACCGGGTCTTGG | 4861 |
| QY | 333 | PheGlyGluGlnPheAlaGlnMetIleGlyIaGluTyrThraLeuValGlnLysSergIy | 352 |
| Db | 4862 | TTCCGCAACAGTTCTGCGGACAGACTCGGTGCGAAAAGTCATGATCAGAAATTCGGAC | 4922 |
| QY | 353 | TyrPheAlaArgAlaSerAlaSerAsnValAspAspMetArgLeuIleLysSergCysAla | 372 |
| Db | 4922 | TACTTCTCCGCTGCGAGCCTCGAAGCAAGCTGACTGAGATCATCGGCGCTGCACT | 4981 |
| QY | 373 | AspLeuAlaValaGluCysAlaPheArgArgJuseGlyValIleGlyHlaAspGluAsp | 392 |
| Db | 4982 | GACCTCGCTGTGCATGTGGCGCTGCGCGGTAAAGATGGCGTTATTCGGCAGAGACAGAG | 5041 |
| QY | 393 | AsnGlyAsnValaIleuArgAlaIleGluPheProArgIleLysGlyLysProPheAsn | 412 |
| Db | 5042 | AACGGCCAGACGTTGACGACATCGCCTTCATCGGATCAAGGGGCGCAAGCCTTTCGAC | 5101 |
| QY | 413 | IleAspThrAspTrpPheAsnSergMetLeuSergIuIleGlyGln----- | 427 |
| Db | 5102 | ACCACTCAGCGGTGATTCACCGCATGCTTCAGCGAGATCGGGCAGGCTGACAGCTATG | 5161 |
| QY | 428 | ProLysGlyGlyLysValGlu | 434 |
| Db | 5162 | CCCAACAGTGGCCCTTCCGAA | 5182 |

| | |
|----|---|
| XX | RESULT 5 |
| XX | ACF64457 |
| ID | ACF64457 standard; DNA; 8530 BP. |
| XX | |
| AC | ACF64457; |
| XX | |
| DT | 17-OCT-2003 (first entry) |
| XX | |
| DE | Propionibacterium acnes DNA contig sequence #23. |
| XX | |
| KM | Acne vulgaris; antisepborrhoeic; dermatological; antibacterial; |
| KW | immunostimulant; immune response; vaccine; ds. |
| XX | |
| OS | Propionibacterium acnes. |
| XX | |
| PN | WO2003033515-A1. |
| XX | |
| PD | 24-APR-2003. |
| XX | |
| PF | 11-OCT-2002; 2002WO-US032727. |
| XX | |
| PR | 15-OCT-2001; 2001US-00978825. |
| XX | |
| PA | (CORI-) CORIXA CORP. |
| XX | |
| PI | Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL; |
| PI | Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D |
| PI | Bartch B, Vallilve-Douglas J; |

XX
DR WPI; 2003-381789/36.
XX
PT New *Propionibacterium* acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a *P. acnes* protein.
XX
XX Claim 1, SEQ ID NO 23, 1481bp; English.
XX

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733) CC
CC encoding a Propionibacterium acnes protein. The invention also relates CC
CC to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to CC
CC immunogenic fragments of P. acnes polypeptides. The invention CC
CC additionally encompasses expression vectors and host cells comprising a CC
CC polynucleotide of the invention; antibodies against polypeptides of the CC
CC invention; fusion proteins comprising a polypeptide of the invention; a CC
CC method for stimulating an immune response specific for a P. acnes CC
CC polypeptide and an isolated T cell population comprising T cells prepared CC
CC via this method; a vaccine composition (comprising P. acnes polypeptides, CC
CC polynucleotides, antibodies, fusion proteins, T cell populations, or CC
CC antigen-presenting cells that express the polypeptide), a method and kit CC
CC for detecting or determining the presence or absence of P. acnes in a CC
CC patient; and a method for inhibiting the development of P. acnes in a CC
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion CC
CC proteins, T cell populations or antigen-presenting cells that express the CC
CC polypeptides are useful for diagnosing, preventing or treating acne CC
CC vulgaris, or for stimulating an immune response specific for a P. acnes CC
CC protein. The polynucleotides can also be used as probes or primers for CC
CC nucleic acid hybridization. The vaccine composition is useful for the CC
CC stimulation of an immune response against P. acnes, or for treating acne, CC
CC and the kit is useful for performing a diagnostic assay. The present CC
CC sequence represents a P. acnes DNA contig which is specifically claimed CC
CC in the invention. Note: The sequence data for this patent did not form CC
CC part of the printed specification, but was obtained in electronic format CC
CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences CC
XQ Sequence 8530 BP, 1921 A, 2633 C, 2423 G, 1551 T, 0 U, 2 Other;

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 2,72e-140 | Length: | 8530 |
| Score: | 1444.50 | Matches: | 274 |
| Percent Similarity: | 80.59% | Conservative: | 54 |
| Best Local Similarity: | 67.32% | Mismatches: | 74 |
| Query Match: | 63.24% | Indels: | 5 |
| DB: | 8 | Gaps: | 1 |

| | | | |
|----|------|--|------|
| QY | 33 | lyelylevvalalalleleutrralaglglyleuulaprocybleuanseralailcelly | 52 |
| Db | 3962 | AMAGAGGTCCGACCTCCACACCGGGGGGAGATTTCGCCGCTACCTATCCACCCGCATCAGT | 4022 |
| QY | 53 | SeurleuilegluhargtryrthrgiulleaspProserilegiulleieCyaryArggly | 72 |
| Db | 4022 | GGATTGATCCAGGCGCTACACCGAGGTAGCACTTAGGCGGAGATCATCGCTTACAGAC | 4082 |
| QY | 73 | GlyTyrIrysglyleuLeuLeuEugIyAspSerTyrProValThrAlaGluValArgIyAsp | 92 |
| Db | 4082 | GGCTACAGAGGGCCCTCGTCAAGGGAGACTTCTCCAGAGTCAACGACACCGTTCGCAAGAT | 4142 |
| QY | 93 | AlaGlyValLeuGlnArgPheGlyGlySerValIleGlyIysAsnArgValIyLeuThr | 112 |
| Db | 4142 | GCCGAGATCCTCAAGCGTTTGGTGATCCCGCATTTGGCAACTGGCGGTCAACCTGACC | 4202 |
| QY | 113 | AsnValIyAspCybValIyAsnGlyLeuValIyGluGlyGluAspProGlnIyVal | 132 |
| Db | 4202 | AACGCCGCCGACCTCGTAAAGCGTGCTGTGCTCCCGAAGCGACGACCCGCTCAAGTT | 4262 |
| QY | 133 | AlaIaAspGlnLeuValIyAspGlyValAspIleLeuIsthrIleGlyIyAspAsp | 152 |
| Db | 4262 | GCCGCCGATCGTCTGGTTCGCCAGCGGGGTGACATTTTGGCACACATCGGTGGTACAC | 4322 |
| QY | 153 | ThrAsnThrAlaAlaIaAspLeuAlaIaPheLeuAlaArgAsnAsnTyrGlyLeuThr | 172 |

| | | | | | | | |
|----------|--|-------------------------|--|------|--|--|--|
| D | 4322 | ACTAATACCA | CGCGCGGCGGATCTGGCGGCTATCTGGCTGAGAAATTAATCAAGGCTCAAC | 4381 | | | |
| Q | 173 | ValIleGlyLeu | ProLYSThrValAspAsnAspValPheProIleLYsGlnSerLeuGly | 192 | | | |
| D | 4382 | GTCTGTGGCC | TGGCGAAGACCATCGATMAACAAGCTCGGATTCGCAATTCCTAGT | 4441 | | | |
| Q | 193 | AlaTrpThrAla | IleGlnGlyValAlaArgYrPheMetAsnValValAlaGlnAsn | 212 | | | |
| D | 4442 | GCTTGAACGG | CTGGCGAACAAGGTTCCGCTTGTCTCAAAATCTGTTGGCGAGACAAAC | 4501 | | | |
| Q | 213 | AlaAsnProArg | MetLeuIleValHISglnValMetGlyArgAsnCyseGlyTrpLeuThr | 232 | | | |
| D | 4502 | TCCGGATCA | CCCAAGCTCATCTGTCACAAATCAATGGCCCTAATCTGGCGTGCCTCAAC | 4561 | | | |
| Q | 233 | AlaAlaThrAla | GlnGlyTyrArgLYsLeuLeuAspArgAlaGluTrpLeuProGlnLeu | 252 | | | |
| D | 4562 | GCCCTACCG | CGCCCAATAACCGGAATGGCTCATACTCAACAGTGCTACAGATC | 4621 | | | |
| Q | 253 | GlyLeuThrArg | GlnSerTyrGlnValHISAlaValPheValProGlnMetAlaIleAsp | 272 | | | |
| D | 4622 | GGGCTGTCCA | GAAGGCTGGAGATCTCACGCCGATATGTCTCCGAAGCTCACATGCAC | 4681 | | | |
| Q | 273 | LeuGlnAla | GlnAlaLYsArgLeuArgGlnValMetAspLYsValAspCyseValAsnIle | 292 | | | |
| D | 4682 | CTTAAGCCGA | AGCCGACCGCTCAACAAGTCAATGACAGAAATTGGCAAGTCACTATC | 4741 | | | |
| Q | 293 | PheValSerGln | GlyAlaGlyValAlaIleValAlaGlnMetGlnAlaLYsGlyGln | 312 | | | |
| D | 4742 | TTCTGTCCGA | GGGGGCTGGCTGAGCCGATTTATGAGAGATGGAAGAACGCTGAG | 4801 | | | |
| Q | 313 | GluValPro | AlaArgAlaPheGlyHISIleLYsLeuAspAlaValAsnProGlyLYsTrp | 332 | | | |
| D | 4802 | GAAGTCCCG | CGCGCATCTCTTGTGCACGTCAAGCTCGCAAGTGAAACCGAGTGTGG | 4861 | | | |
| Q | 333 | PheGlyGln | GlnPheAlaGlnMetIleGlyAlaGlyLYsTrpLeuValGlnLYsSerGly | 352 | | | |
| D | 4862 | TTCCGCAAC | AGTTGCGGACAAGCTCGGATCCGAAGAGTCAATGTTCAAGATCCGAC | 4921 | | | |
| Q | 353 | TyrPheAla | ArgAlaSerAlaSerAsnValAspAspMetArgLeuIleLYsSerCyseAla | 372 | | | |
| D | 4922 | TACTTCTCC | GCTGGAGCTCGAACAGACTGACCTGAGGCTCATCTGGCGGCTGACT | 4981 | | | |
| Q | 373 | AspLeuAla | ValGlyCyseAlaPheArgArgLYsSerGlyValIleGlyHISAspGlnAsp | 392 | | | |
| D | 4982 | GACCTCGGT | GCATGTGGCGCTGCGCGCTGAGCTGCTTATCTGCGCACAGACGAGAG | 5041 | | | |
| Q | 393 | AsnGlyAsn | ValLeuArgAlaIleGlnPheProArgIleLYsGlyLYsArgProPheAsn | 412 | | | |
| D | 5042 | AACGGCGAC | AGCTTGACGACATGCTTCGATCGGATCAAGGGCGGCAACCTTTGCAC | 5101 | | | |
| Q | 413 | IleAspThr | AspTrpPheAsnSerMetLeuSerGlnIleGlyGln----- | 427 | | | |
| D | 5102 | ACCACTCA | CGCGGTGTTACACCGCATCTGACGCGAGATCGGCGAGGCTGACAGTATGG | 5161 | | | |
| Q | 428 | ProLYsGly | GlyLYsValGln 434 | | | | |
| D | 5162 | CCCAACAG | TGGCCCTTCGAA 5182 | | | | |
| RESULT 6 | | | | | | | |
| ID | ADS62361 | | | | | | |
| TD | ADS62361 | standard; cDNA; 552 BP. | | | | | |
| XX | ADS62361; | | | | | | |
| XX | AC | | | | | | |
| XX | 02-DEC-2004 | (first entry) | | | | | |
| DS | Bacterial polynucleotide #14348. | | | | | | |
| KM | Recombinant DNA construct; transformed plant; improved plant property; | | | | | | |
| KM | cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic; | | | | | | |
| KM | pathogen tolerance; pest tolerance; plant disease resistance; | | | | | | |
| KM | cell cycle pathway modification; plant growth regulator; | | | | | | |

homologous recombination; seed oil yield; protein yield; carbohydrate;
 nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 bacterial polynucleotide; gene; ss.

Bacteria.

US2003233675-A1.

18-DEC-2003.

20-FEB-2003; 2003US-00369493.

21-FEB-2002; 2002US-0360039P.

CAO Y, HINKLE GJ, SLATER SC, CHEN X, GOLDMAN BS;
 WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide
 for expression of a polynucleotide encoding a polypeptide from a
 microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 38035; 122pp; English.

The invention relates to a recombinant DNA construct comprising a
 promoter functional in a plant cell, where the promoter is positioned to
 provide for expression of a polynucleotide encoding a polypeptide from a
 microbial source. The invention also relates to a transformed plant
 comprising the recombinant DNA construct and a method of producing a
 transformed plant having an improved property. The plant is a crop plant
 such as maize or soybean. The method of producing a transformed plant
 having an improved property comprises transforming a plant with the
 recombinant DNA construct and growing the transformed plant, where the
 polynucleotide or polypeptide is useful for improving plant properties.
 The recombinant DNA construct is useful for producing plants with
 improved plant properties, e.g. improved cold, heat or drought tolerance,
 tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 increased resistance to plant disease, better growth rate by modification
 of the cell cycle pathway with plant growth regulators, increased rate of
 homologous recombination, modified seed oil or protein yield and/or
 content, improved yield by modification of carbohydrate, nitrogen or
 phosphorus use and/or uptake, by modification of photosynthesis or by
 providing improved plant growth and development under at least one stress
 condition. Improved lignin production or improved galactomannan
 production. This sequence represents a bacterial polynucleotide used in
 the scope of the invention. Note: The sequence data for this patent did
 not form part of the printed specification but was obtained in electronic
 format from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 552 BP, 109 A, 181 C, 166 G, 96 T, 0 U, 0 Other;

Alignment Scores:

| Pred. No.: | 1,63e-59 | Length: | 552 |
|------------------------|----------|---------------|-----|
| Score: | 659.00 | Matches: | 125 |
| Percent Similarity: | 82.61% | Conservative: | 27 |
| Best Local Similarity: | 67.93% | Mismatches: | 32 |
| Query Match: | 28.85% | Indels: | 0 |
| DB: | 13 | Gaps: | 0 |

US-10-701-200-6 (1-437) x ADS62361 (1-552)

33 LyslySVALAIAIleLeuThRAIaglyGlyLeuAlaProCYLeuAbnSeraIaileGly 52
 ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 1 CAGAAAGTCGAATGCTGACCGCGGGTGGCTCGCGCCCTGCTCTTCCTCGGCGCGGGT 60
 53 SerleuIleGluwAGTGTThcdulLeaapProserIleGluIleleCYerYrargGly 72
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

| | | | |
|----------|---|--|-----|
| Db | 61 | GGCCGTGATGACGCGCTATGACGACATCGCGCCGTAATGACATCATCGCTTACCGCTCC | 120 |
| Qy | 73 | GLYTRVLSGLYLEULEULEUGLYASPSERTYRPROVALTHRLAAGLVALAAGLYSLYS | 92 |
| Db | 121 | GGTTATACAGGGGGCTGCTTGTGGCGGCGGTATCGAGATCACCAAGATATGCGCGAAAG | 180 |
| Qy | 93 | ALAAGLYVALLEUGLNAKRPHEGLYGLYSERVALILEGLYASNSERARGLVALYSLEUTH | 112 |
| Db | 181 | GGGCATCTACTCATCTGTTATACGGCGGCTGCGCATGGCAACAGCCGCGTCAAGCTGACC | 240 |
| Qy | 113 | AAenValLYsASPcYservalLYsARGLYsLeuVALYsGLYsGLYsAspProGLYsVAL | 132 |
| Db | 241 | AATGCCGCCGATTTGCGCCAAAGCGCGGCTCGTCAAGAAAGGCGACAATCCGTTGCGCGTC | 300 |
| Qy | 133 | ALAALAAspGLINLeuVALYsAspGLYsVALAspPILeUHIsthrILEGLYsGLYsAsp | 152 |
| Db | 301 | GGCGGTGAACGGCTGGCGGCTGACCGGCATCACCATTTCTCCATCTATCGGTGCGACGAC | 360 |
| Qy | 153 | ThraenThrALAALAAspLeuVALAspLeuVALAspLeuVALAspAsnTYRGLYsLeuThr | 172 |
| Db | 361 | ACCAACACCAACCGCGCGGACCTTGGCGCTTATCTGCGGCCAACGAGATATGACTTCACC | 420 |
| Qy | 173 | VALILEGLYLeuProLYsThrVALAspAsnAspVALAspProILEYsGLINserLeuGLY | 192 |
| Db | 421 | GTCTGTGGGCTGCGCTGCGAAGACGGGTGATTAATGACGTGTGCGCATCAAGCAGTGTGCGC | 480 |
| Qy | 193 | ALAtrPTHrALALAAGLUGLNGLYALAARGTYRPhemeCAsnVALAVALAAGLUAAsn | 212 |
| Db | 481 | GCCTGAGACGGCCGCGAAGTCGGTGTGCTGCTTCTTGCAACATGTGCAACAGCAGCAGAC | 540 |
| Qy | 213 | ALAAsnProARG 216 | |
| Db | 541 | GGCGCACCGAAA 552 | |
| RESULT 7 | | | |
| ADS56121 | | | |
| ID | ADS56121 | standard; cDNA; 1050 BP. | |
| XX | AC | ADS56121; | |
| XX | DT | 02-DEC-2004 (first entry) | |
| XX | DE | Bacterial polynucleotide #8108. | |
| XX | XX | Recombinant DNA construct; transformed plant; improved plant property; | |
| XX | KM | cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; | |
| XX | KM | pathogen tolerance; pest tolerance; plant disease resistance; | |
| XX | KM | cell cycle pathway modification; plant growth regulator; | |
| XX | KM | homologous recombination; seed oil yield; protein yield; carbohydrate; | |
| XX | KM | nitrogen; phosphorus; photosynthesis; lignin; galactomannan; | |
| XX | OS | bacterial polynucleotide; gene; ss. | |
| XX | OS | Bacteria. | |
| XX | XX | US2003233675-A1. | |
| XX | PN | 18-DEC-2003. | |
| XX | PD | 20-FEB-2003; 2003US-00369493. | |
| XX | PF | 21-FEB-2002; 2002US-0360039P. | |
| XX | PR | | |
| XX | XX | (CAOY/) CAO Y. | |
| XX | PA | (HINK/) HINKLE G J. | |
| XX | PA | (SLAT/) SLATER S C. | |
| XX | PA | (CHEN/) CHEN X. | |
| XX | PA | (GOLD/) GOLDMAN B S. | |
| XX | PI | Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS; | |
| XX | XX | WPI, 2004-061375/06. | |
| XX | XX | | |
| PT | New recombinant DNA construct comprising a promoter positioned to provide | | |

Claim 1, SEQ ID NO 31795; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomanan
CC production. This sequence represents a bacterial plasmid construct used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seeddata.uspto.gov/sequence.html.

SQ Sequence 1050 BP; 187 A; 374 C; 340 G; 149 T; 0 U; 0 Other;

| | |
|------------------------|----------|
| Alignment Scores: | |
| Pred. No.: | 3,76e-22 |
| Score: | 306.50 |
| Percent Similarity: | 42.49% |
| Best Local Similarity: | 26.46% |
| Query Match: | 13.42% |
| DB: | 13 |
| | |
| Length: | 1050 |
| Matches: | 104 |
| Conservative: | 63 |
| Mismatches: | 145 |
| Indels: | 81 |
| Gaps: | 14 |

US-10-701-200-6 (1-437) X ADS56121 (1-1050)

| | | | | |
|----|--|-----|---|-----|
| OY | | 23 | PheAsnPhetYrThrIleMetCasnlysrProlysylsvAlaIleuethrAlaGly | 42 |
| Dd | | 10 | TTCAAGAGATACGCCAAT-----CGATCGGGGTACTGCACGGGGTGGC | 57 |
| OY | | 43 | LeuAlaProCybLeuAsnSerAlaIleGlySerIleuIleGlArgyTrhrgluIleAp | 62 |
| Dd | | 58 | GACAGCCCCGGGCGCTGAACGGCGGTATCCGTGCGGTGTCGTAAGGCATCAAG---GAA | 114 |
| OY | | 63 | ProserIlegluuIleIleCyetrArgglygIyTrLyvsglyLeuleuLeuglyAapSer | 82 |
| Dd | | 115 | TACGGATTACGATTGTCGTGGAATTCCTGTACGGGTGGCGGGCCCCCTTGAAAGCCACACC | 174 |
| OY | | 83 | TyrProValThrAlaGluValArglybLybAlaGlyValLeuGlnArpHehGlyGlySer | 102 |
| Dd | | 175 | ATGCCCGCTGCACATCGAAGCGGTCCG-----GGCATCTCTCCGGC---GGCGGTACC | 225 |
| OY | | 103 | ValIleGlyAsnSerArgVallybLeuthrAsnValLybAPCyavaLybsArglyLeu | 122 |
| Dd | | 226 | ATCTCGGCTCTCCGCGCACCAACTCAATGAAGATCGAGGGCGGTGTCCAGCGG----- | 279 |
| OY | | 123 | VallybgluGlylAapProGlnlybValAlaIleApGlnLeuValLybsApGlyVal | 142 |
| Dd | | 280 | -----GTCAAGACAACATGGCCGGCGCTCGGGGTG | 309 |
| OY | | 143 | AspIleLeuHrsthrlleGlyLybAPsrThraenthralAlaAlaAspLeuAlaIla | 162 |
| Dd | | 310 | GACGGCGTCTGTCGCAATCGCGGGGTAGAAGACACCTTCGGCGTGGCGGCCACAGCTC----- | 363 |
| OY | | 163 | PheleuAlaArgAsnAntyGlyLeuthrValIleGlyLeuProlysrThrValLaapAn | 182 |
| Dd | | 364 | -----CAGACACACGGTGTCAACGTGGTGGGTGGTGGCCCAAGACCATGACCAAC | 411 |

PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144681P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158025P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158393P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159322P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 2,47e-19 Length: 1458
Score: 281.50 Matches: 109
Percent Similarity: 41.56% Conservative: 61
Best Local Similarity: 26.65% Mismatch: 166
Query Match: 12.32% Indels: 73
Gaps: 16

US-10-701-200-6 (1-437) x AAC46084 (1-1458)

QY 37 l l e u t h r a l a g l y l e u d l a p r o c y s l e u s e n s e r a l l i e g l y s e r l e u l l e c t u 56
DB 286 A T A G T T A C A T G T G G A G C T C T G T C C C G T C C A T A C C G T T A G A G A A C T T G A C 345
QY 57 A r g T y r T h r G l u i l l e a s p p r o s e r l l e g l u l l e i l e c y t y z a r g g l y g l y t y r y s g l y 76
DB 346 A G C T T A T C A T A C A T A C T A T G A G A A G A A T T C T T G A A T C G A T G G A T A C A G A G A 405
QY 77 l e u l e u l e u g l y a s p s e r t y r p r o v a l t h r a l a g l u a l a r g l y a l a g l y v a l l e u 96
DB 406 T T T T A C G C C A G A A A T A C A T C C C T T A A C T T A A G T C ----- G T G A A T G A T A T C 456
QY 97 G l n a r g p h e g l y l y s e r v a l l i e g l y a n s e r a r g v a l l y s l e u t h r a s n v a l l y a s p 116
DB 457 C A T A A C C G A G A G A A C A T C A T T G G A C C T C A A ----- 492
QY 117 C y s v a l l y s a r g g l y l e u v a l l y s g l u g l u a s p p r o g l i n l y e v a l a l a l a s p g l n 136
DB 493 ----- G T G A C A C T G A T A C C A C A A G A T A --- G T T A T A G C 525
QY 137 l e u v a l l y a s p g l y a l a s p i l e u t h r a l a g l y l e u l l e c y l y a s p a s p t h r a s n t h r a l a 156
DB 526 A T T C A A G T A C G A G A A T A C A T C A G G T T A C A T T A T T G A G A G A T G A A C C A G A G A G T 585
QY 157 A l a l a l a s p l e u a l a l a l a p h e u l a l a r g a s n e n t y r g l y l e u t h r a l l e g l y l e u 176
DB 586 G C T T C A G T A A T A T T T G A G A A T T A G A A G C T C A T A A A G T C T G C T G T G G A A T T 645
QY 177 p r o l y s t h r a l a s p a s p v a l p h e p r o l l e y s g i n s e r l e u g l y a l a t p h r a l a 196
|||||

| | | | | | | | | | | | | |
|----|------|------------------------|-----------|-----------|--------------|-------------|------------|-------------|-------------|-----------|---------|------|
| Oy | 74 | TYRISGLILEUleuLeuGLIYA | SPSERTYR | PROVAL | THIRL | AGL | VAL | ARG | LY | VAL | 93 | |
| Db | 361 | TATCAAGG | CTTCTTCC | TGAGTCAAA | TTCAC | CTTTATTAAT | CTTAATCC | AGATG | TTGTT | | 420 | |
| Oy | 94 | GLI | VAL | LEU | GIN | ARG | PH | GLI | Y | GLI | Y | 113 |
| Db | 421 | GATGATATTA | TAAATTTGG | AGTACTAT | TCCTGGTCTTCA | ----- | | | | | 462 | |
| Oy | 114 | VAL | YS | AS | PC | Y | VAL | LY | ARG | GLI | Y | 133 |
| Db | 463 | ----- | | | | AGGGG | TGCTATTAAG | CCCTGGAA | ----- | ATA | 489 | |
| Oy | 134 | ALA | AS | RG | IN | LEU | VAL | LY | AS | PG | I | 153 |
| Db | 490 | GTTGAT | TACTTTAG | AGAA | TGA | CACTTAAT | TGATTTTA | CACTGGG | GGGATG | TA | 549 | |
| Oy | 154 | AS | TH | IR | LA | LA | LA | AS | PR | LEU | LA | 173 |
| Db | 550 | CAAAAGG | CTCTCTTAT | TGCTG | AGAGATG | AAAAA | CAATTTAAAA | TATG | ACACT | | 609 | |
| Oy | 174 | IL | GLI | LEU | P | CO | LY | TH | VAL | AS | PA | 193 |
| Db | 610 | GTTGG | CACTCTTA | AAACCT | TAG | CAATATTTTAT | TGTTGTTCC | AAAAATCTTTG | GGATTT | | 669 | |
| Oy | 194 | TRP | TH | IR | LA | LA | AG | IN | GLI | Y | LA | 213 |
| Db | 670 | GAG | ACTGCTG | TAGA | AA | CAAGCTG | TGCAC | CTGTGCTG | TCATTTGA | AGCTAAT | 729 | |
| Oy | 214 | AS | N | PRO | AS | ME | LEU | IL | LEU | LA | 233 | |
| Db | 730 | GCTTA | TAAATG | GCATTTG | ACCTTTAA | GTATTTG | GAACGG | AAATCTTG | CTTTATTTG | CTCT | 789 | |
| Oy | 234 | ALA | TH | IR | LA | AG | IN | GLI | Y | TR | ARG | 253 |
| Db | 790 | CATACCGA | ----- | | | | | | | | 798 | |
| Oy | 254 | LEU | TH | IR | ARG | IN | SE | TYR | GLI | Y | LA | 273 |
| Db | 799 | ----- | | | | CTTTCTTCA | TGATGTTAT | TTTGTGTTAA | TCCAGAA | CTTGATTTG | 852 | |
| Oy | 274 | GLI | AL | LEU | AG | IN | ----- | | | | 284 | |
| Db | 853 | GAA | GGTCTTA | TGATTTCTT | GTTCATCTT | GAA | AGAC | CACTTTAG | AAAA | AGAA | 912 | |
| Oy | 285 | AS | PI | Y | VAL | LA | AS | PI | Y | VAL | LA | 304 |
| Db | 913 | GAA | AGATTTCC | CCCA | TGAC | TAATATTTG | TAC | AGAA | GAG | AC | 957 | |
| Oy | 305 | ALA | GLI | ME | T | GI | LA | LY | GLI | Y | LA | 322 |
| Db | 958 | ----- | | | | CAAAA | TATTTTGAT | TCATTTTCC | TAA | GA | 1008 | |
| Oy | 323 | LYS | LEU | AS | PA | LA | VAL | AS | PA | LA | 342 | |
| Db | 1009 | CTTAT | TATGAG | ATAT | ----- | GGGCTT | TATATTA | AAAGATAA | ATATAC | AGATATTT | 1065 | |
| Oy | 343 | ALA | GLI | Y | ----- | | | | | | 358 | |
| Db | 1063 | GCA | AAATAT | TACA | TTTACTT | TAAAT | TGATTC | CCAC | ACTATATTTAT | TAGAA | 1122 | |
| Oy | 359 | ALA | SE | PA | EN | VAL | AS | PA | ME | C | ARG | 378 |
| Db | 1123 | CTG | CTAAT | TG | CCAG | TATCG | CTTATTTG | CTG | TAG | CTTG | GGATG | 1182 |
| Oy | 379 | ALA | PH | EAR | GA | RG | IN | SE | TYR | GLI | Y | 398 |
| Db | 1183 | GCA | ATG | CGCTG | TGA | AGCA | AAATATG | TGATTTAG | TATG | AGATAC | AAATTTG | 1242 |
| Oy | 399 | ALA | IL | GLI | Y | PH | EAR | GA | RG | IN | SE | 418 |
| Db | 1243 | CC | TA | TAT | AGATG | GC | AG | TATATTA | G | ACAA | ATATAG | 1299 |

| | | | |
|----|--|--|------------------|
| 0y | 411 | AmsSerMeLeSerCluIleGlyClnPro | 428 |
| | | | |
| | | | |
| Db | 1300 | AGAGATGTTCTTCAAGTACGGGGCAACCA | 1329 |
| | | | |
| | RESULT 10 | | |
| ID | AA085983 | | |
| XX | AA085983 | standard; cDNA to mRNA; 1778 BP. | |
| XX | | | |
| XX | AA085983; | | |
| XX | | | |
| DT | 25-MAR-2003 | (revised) | |
| DT | 11-OCT-1995 | (first entry) | |
| XX | | | |
| DE | Flaveria brownii PFK-FBI gene. | | |
| XX | | | |
| KW | ATP-dependent fructose-6-phosphate 1-phosphotransferase enzyme; plant; | | |
| KW | potato; Solanum tuberosum; rice; Oryza sativa; maize; Zea mays; radish; | | |
| KW | Raphanus sativus; Flaveria brownii; primer; expression vector; | | |
| KW | Agrobacterium tumefaciens; sugar; storage; temperature; ds. | | |
| XX | | | |
| OS | Flaveria brownii. | | |
| XX | | | |
| FT | Key | Location/Qualifiers | |
| FT | CDS | 19..1470 | |
| FT | | /tag= a | |
| FT | | /product= "fructose-6-phosphate 1-phosphotransferase" | |
| XX | | | |
| PN | W09505457-A1. | | |
| XX | | | |
| PD | 23-FEB-1995. | | |
| XX | | | |
| PF | 16-AUG-1994; | 94MO-JP001352. | |
| XX | | | |
| PR | 19-AUG-1993; | 93JP-00226454. | |
| XX | | | |
| PA | (NTSB) JAPAN TOBACCO INC. | | |
| XX | | | |
| PI | Hiyoshi T, Mine T, Kaseoka K, Tyson HR, Page MJA; | | |
| DR | WPI; 1995-098757/13. | | |
| XX | | | |
| DR | P-PSDB; AAR71580. | | |
| XX | | | |
| PT | DNA coding for fructose-6-phosphate 1-phosphotransferase - of plant | | |
| PT | origin, for prodn. of transformant plant cells with altered sugar | | |
| XX | content. | | |
| XX | | | |
| PS | Claim 6; Page 39-43; 79pp; Japanese. | | |
| XX | | | |
| CC | The sequences (AA085982-6) represent the genes encoding a novel ATP- | | |
| CC | dependent fructose-6-phosphate 1-phosphotransferase enzyme (EC 2.7.1.11; | | |
| CC | PFK) from a range of plants. This sequence is the Flaveria brownii gene | | |
| CC | PFK-FBI. Plants transformed with these genes can express the enzyme. The | | |
| CC | transformed plants can produce varieties that have altered sugar content | | |
| CC | on storage at low temperatures. (Updated on 25-MAR-2003 to correct PN | | |
| XX | field.) | | |
| XX | | | |
| SQ | Sequence 1778 BP; 521 A; 312 C; 420 G; 525 T; 0 U; 0 Other; | | |
| | | | |
| | Alignment Scores: | | |
| | Pred. No.: | 5, 22e-18 | Length: 1778 |
| | Score: | 270.00 | Matches: 113 |
| | Percent Similarity: | 40.59% | Conservative: 53 |
| | Best Local Similarity: | 27.63% | Mismatch: 175 |
| | Query Match: | 11.82% | Indels: 68 |
| | DB: | 2 | Gaps: 15 |
| | | | |
| | US-10-701-200-6 (1-437) x AA085983 (1-1778) | | |
| 0y | 36 | AlAlleLeuThrAlaGlyGlyLeuAlaProCysLeuAenSerAlaIleGlySerIleuLe | 55 |
| | | | |
| | | | |
| Db | 301 | GCAATAGTAAACATGCGTGTATTATGCTCTGGCGCTAAACACAGCATCAGGAAATGTT | 360 |
| | | | |
| 0y | 56 | GluaGlyTyrThrClnuLeaAspProSerIleGluIleIleCysTyrArgGlyTyrIle | 75 |

| | | | |
|----|------|---|------|
| Db | 361 | TCGGCAGCTTTATCACATGATATGGTGACCAAAAGTTCTTGGAGATTGATGGAGGTCACGA | 420 |
| Oy | 76 | leuleuleu gl y a ser p ro val thr a gl val a gl y a a gl y a | 95 |
| Db | 421 | GGTTTTCACAAAAAACACCATCAGCTTGTGACTCCAAAGGTT-----GTGAATGAC | 471 |
| Oy | 96 | leu g n a g p he g l y s e r v a l l e g y a n s e r a g v a l l y e u t h r a n v a l l y a | 115 |
| Db | 472 | ATTCATAAACCGTGGGTGATCAATTAATGGCAGCTTCCTGCT----- | 510 |
| Oy | 116 | a s p c y s v a l y a s g l y e u v a l l y s g l y g l a n p p r o g l n y e v a l a a a s p | 135 |
| Db | 511 | -----CGGGGACATGATTAACCAAAATAT--GTTGAC | 540 |
| Oy | 136 | g l n e u v a l y a s p g l y a a s p l l e u n i s t h r i l e g l y g l a s p s t h r a n t h r | 155 |
| Db | 541 | AGTATTCAGATCCGTGATCATCATAGGTTATATTAATTGGAGAGAGACGATCTCAAAAG | 600 |
| Oy | 156 | l a a l a a a s p l e u a a l a p h e u a l a a g b a n a n t y g l y e u t h r v a l l e g l y | 175 |
| Db | 601 | CGAGCAGCTGTTTATTCAGGAAGTGGAAGGCGTGCGCTTAAAGCTGTAGTGCGTGGG | 660 |
| Oy | 176 | l e u p o l y e t h r v a l a s p a s p v a l a p e r o l l e y s g l n s e r l e u g l y a t h r p h r | 195 |
| Db | 661 | ATTCCTTAAGACAATTGATTAATGACATTCGCGCATTTGATTAAGCTCTTTGGTTTGACACG | 720 |
| Oy | 196 | l a a l a g l n g l y a a a g t y r p h e t e r a n v a l a l a g l u a n s a n a a s p r o | 215 |
| Db | 721 | GCTGTGGAGAGAGGCTCAACGTCGCCCTTAATGCTGCACATGTGGAGCGTGAAAGTCTGAG | 780 |
| Oy | 216 | a r g w e t e u i e v a i n s g l u a l w e g l y a n g b a n c y s g l y t h r p l e u t h r a a l a t h r | 235 |
| Db | 781 | AATGCGATAGAGGGGTGCTCAAACTTATGTGGAGCGCTATGATGATTCAGCA----- | 831 |
| Oy | 236 | a l a g i n l u y r a g l y e u l e u a s p a g l u t h r p l e u p r o g i n l e u g l y e u t h r | 255 |
| Db | 832 | -----ATGTATGCACT | 843 |
| Oy | 256 | a r g l s e r t y r g l u v a i n a l a v a l p h e v a l p r o g i u m e t a i l a s p l e u g l u a | 275 |
| Db | 844 | TTGCGTAGTCGAGATGTTGATTTATGTTAAATCTGTAATCACCTTTTATCTTGAGGGA | 903 |
| Oy | 276 | g l u a a -----l y a a g l e u a a g l u a l w e t a s p l y e v a a s p | 288 |
| Db | 904 | GAAAGTGACCTTTAGATATGTAGTGAAGAAAAGCTGTCAAGAGCAATGACACATGCTC--- | 960 |
| Oy | 289 | c y s v a l a n i l p h e v a l s e r g l u g l y a g l y a l g l u a l l e v a i l g l u w e g i n | 308 |
| Db | 961 | -----ATCGTTGTGACGAAGAGCTGCTGTCAGAGCTGCTTGACGACGAAGAAACCTTG | 1011 |
| Oy | 309 | a l a l y s g l n g i n v a l p r o a g b e r a l p h e g i y i s i l y e v e u a s p a l e v a l n | 328 |
| Db | 1012 | AAA-----ACCTCAACCGCAAAAGATCTTCTGGAATTAACATTACTTCACGATGTC--- | 1062 |
| Oy | 329 | p r o g i l y e t h r p h e g i l y g l n -----p h e a l g i n e i l e g l y a l a g l u | 344 |
| Db | 1063 | ---GGATTGTGATTTCTGATAGATTAAAGCTCACCTTGTGTAATAATCTCCCATGCT | 1119 |
| Oy | 345 | l y e t h r e u v a l n -----l y s e r g l y r p h e a l a g l a s e r a l s e r a n v a a s p | 363 |
| Db | 1120 | ATTACTCTCAATACATAGATCSAACTTAATGATCCGTGCGGTTCCAAAGTAATGATCATCT | 1179 |
| Oy | 364 | a s p e r e t a s g l e i l y e s e r c y a l a s p l e i u a l g i u c y a l p h e a g a r g s i u | 383 |
| Db | 1180 | GATAAT-----GTATCTGCACTCTCTTGCTCAAAAGTTGTTCATGAGAGTGATG | 1230 |
| Oy | 384 | s e r g i y a l l e g l y i a s p l u a p -----a e n g i y a n v a i l e u a g l a l l e g i u | 401 |
| Db | 1231 | CGGGGCTACACCGGCTTCAAAAGTGGGCTTGCAATGATGA---CAAGATTAATTCGA | 1287 |
| Oy | 402 | p h e r o a g i l l y e g l y l y e b r o h e a i n i l e a r t h r a p -----t h r p h e i n | 419 |

[illegible]

[illegible]

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Kategiri F, Quan S, Tao Y, Whitcham S, Xie Z, Zhu T, Zou G;
PI WPI, 2003-175290/17.

DR Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.

PS Claim 6; SEQ ID NO 3978; 899bp; English.

XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.

XX Sequence 2256 BP; 568 A; 574 C; 665 G; 449 T; 0 U; 0 Other;

SO Alignment Scores:

| Pred. No.: | 3.47e-17 | Length: | 2256 |
|------------------------|----------|---------------|------|
| Score: | 263.50 | Matches: | 96 |
| Percent Similarity: | 40.71% | Conservative: | 53 |
| Best Local Similarity: | 26.23% | Mismatches: | 158 |
| Query Match: | 11.54% | Indels: | 59 |
| DB: | 8 | Gaps: | 9 |

US-10-701-200-6 (1-437) x ADA70655 (1-2256)

QY 34 LysValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53

Db 487 AAGCATGCATTGTAACTGCGCGAGCCCTTCCCTGGCGTAAATCGGTCAATGAGAG 546

QY 54 LeuIleGluArgTyrThrCulIleAspProSerIleGluIleIleCysTyrArgGlyGly 73

Db 547 TTGGTGGCGCGCTTCCACATGTAAATGCAATGATATCTTCGGCATACGAATGCA 606

QY 74 TyrIleGlyLeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgIleValAla 93

Db 607 TACAAGGAGATTCTATTCAGATTAATATCTCTCTATGACACCT-----AAAGTGC 657

QY 94 GlyValLeuGluArgPheGlyGlySerValIleGlyAsnSerArgValIleLeuThrAsn 113

Db 658 AATGATATCCCAAAAAGAGGTGGACAGTCTGGGAACATCCAG-----702

QY 114 ValIleAspCysValIleAspArgIleValIleValIleGluGlyAspProGlnIleValAla 133

Db 703 -----GGTGGCATGACCCAAAAAAT--726

QY 134 AlaAspGlnLeuValIleAspGlyValAspIleLeuHisThrIleGlyGlyAspArgThr 153

Db 727 GTTGCAACATTCGACATCGTGTATTAATCAAGTATATCATTCATTCGAGCGATGGAAC 766

QY 154 AsnThrAlaAlaIleAspIleValAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrVal 173

Db 787 CAGAAGGAGCATATGAATATATATGAAGAAATTCGACAGCTGAGCTAAAGTTGCTGT 846

QY 174 IleGlyLeuProIleValIleAspAsnAspValPheProIleIleGlnIleLeuGlyAla 193

Db 847 GCTGGTGTTCCTCCAAACAATGATATATATTAAGCGGTATATGACAAAGTCTTTGGT 906

QY 194 TrpThrAlaAlaGluGlnGlyValIleArgTyrPheMetAsnValIleAlaGluAsnAsnAla 213

Db 907 GATAGTGTCTGAAGAAGAGCCAGCGTGCATGATGACAGCTCATCTTGAAGCTTCAAG 966

QY 214 AsnProArgMetLeuIleValHisGluValMetGlyArgAsnCysGlyTrpLeuThrAla 233

| | | | |
|-----------|--|--|------|
| Qy | 257 | GIuSeTYrGluValAlhAlaValAlPheValProGluMetAlaIleAspLeuGIuValGlu | 276 |
| Db | 700 | -----GACGTGACCTGCTGCTATCCCGGAAATCTCATTTTCTTGAAGGCAAA | 750 |
| Qy | 277 | -----AlaLysArgLeuArgGluValMetAspLysValAspCys | 289 |
| Db | 751 | GGCGGCTCTTTGCACTTATTCGGTAAACGGCTAAAGAGATTGGTCACATGGT----- | 804 |
| Qy | 290 | ValAsnIlePheValSerGIuGIyAlaGIyValGIuAlaIleValAlaGIuMetGlnAla | 309 |
| Db | 805 | -----ATTGATGATAGACAGAAAGGTGCTGA---CAAGATCTGTGGCTGGAAGCAATGA | 855 |
| Qy | 310 | LyseGIyGlnGIuValProArgAspAlaPheGlyAlaIleLysLysLeuAspAlaValAsnPro | 329 |
| Db | 856 | CAGTCCACMAACCTTC---AAAGATGTCATCTGGGAACAAATCTTACMAAGAGCTT----- | 906 |
| Qy | 330 | GIyLysTPheGluGIuGlnProPheAlaGlnMetIleGIyAlaGIuLysThrLeuValGln | 349 |
| Db | 907 | GGCCTATGATGATCTCCCAACGGATCTCAAGATCATTTTGGCCAAAGATGATACCTTAAACCTG | 966 |
| Qy | 350 | LYS-----SerGIyTYrPheAlaArgAlaSerAlaSerAsnValAspAspMetArg | 366 |
| Db | 967 | AAATATCATAGATCCACACTCATGATATAGGGCTGTCCGACCATATGCATCAGAC----- | 1020 |
| Qy | 367 | LeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArgGluSerGIyAla | 386 |
| Db | 1021 | -----AATGTATGCTGCACCGCTGTAGCTCAAAAGCCGGTTCAT-----GGAAGTG | 1065 |
| Qy | 387 | Ile---GlyAlaProGluAspAsnGlyAsnValIleuAlaGlnA-----IleGlu | 401 |
| Db | 1066 | ATGGCTGATACAAATGGCTTCAACCGTGGCTCTGTAAATGCGACATATCTTACATTTCC | 1125 |
| Qy | 402 | PheProArgIle-----LysGIyGIyLysProPheAsnIleAspThrAspTrPheAsn | 419 |
| Db | 1126 | TTCTATAGATCATCTGAGAAACAGAAACAGAGGTGATCAGTACAGATGTGGCAAGG | 1185 |
| Qy | 420 | SerMetLeuSerGIuIleGIyGlnProLysGIyLysValGIuValSerHis | 437 |
| Db | 1186 | ---CTTTGTCTTCGACAAACAGCGCTAGTTTCATGAAGCAGATGATCACCAC | 1236 |
| RESULT 15 | | | |
| AD848162 | | | |
| ID | AD848162 | standard; cDNA; 960 BP. | |
| XX | AD848162; | | |
| XX | 02-DEC-2004 | (first entry) | |
| XX | Bacterial polynucleotide #2905. | | |
| XX | Recombinant DNA construct; transformed plant; improved plant property; | | |
| KW | cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; | | |
| KW | pathogen tolerance; pest tolerance; plant disease resistance; | | |
| KW | cell cycle pathway modification; plant growth regulator; | | |
| KW | homologous recombination; seed oil yield; protein yield; carbohydrate; | | |
| KW | nitrogen; phosphorus; photosynthesis; lignin; galactomannan; | | |
| KW | bacterial polynucleotide; gene; ss. | | |
| XX | Bacteria. | | |
| OS | Bacteria. | | |
| XX | US2003233675-A1. | | |
| XX | 18-DEC-2003. | | |
| XX | 20-FEB-2003; 2003US-00369493. | | |
| PF | 21-FEB-2002; 2002US-0360039P. | | |
| PR | | | |
| XX | (CAOY/) CAO Y. | | |
| PA | (HINK/) HINKLE G J. | | |
| PA | (SLAT/) SLATER S C. | | |
| PA | (CHEN/) CHEN X. | | |
| PA | (GOLD/) GOLDMAN B S. | | |

| | |
|----|---|
| XX | Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS; |
| PI | WPI; 2004-061375/06. |
| DR | |
| XX | |
| PT | New recombinant DNA construct comprising a promoter positioned to provide |
| PT | for expression of a polynucleotide encoding a polypeptide from a |
| PT | microbial source, useful for producing plants with improved properties. |
| XX | |
| P5 | Claim 1; SEQ ID NO 26592; 122pp; English. |
| XX | |
| CC | The invention relates to a recombinant DNA construct comprising a |
| CC | promoter functional in a plant cell, where the promoter is positioned to |
| CC | provide for expression of a polynucleotide encoding a polypeptide from a |
| CC | microbial source. The invention also relates to a transformed plant |
| CC | comprising the recombinant DNA construct and a method of producing a |
| CC | transformed plant having an improved property. The plant is a crop plant |
| CC | such as maize or soybean. The method of producing a transformed plant |
| CC | having an improved property comprises transforming a plant with the |
| CC | recombinant DNA construct and growing the transformed plant, where the |
| CC | polynucleotide or polypeptide is useful for improving plant properties. |
| CC | The recombinant DNA construct is useful for producing plants with |
| CC | improved plant properties, e.g. improved cold, heat or drought tolerance, |
| CC | tolerance to herbicides, extreme osmotic conditions, pathogens or pests, |
| CC | increased resistance to plant disease, better growth rate by modification |
| CC | of the cell cycle pathway with plant growth regulators, increased rate of |
| CC | homologous recombination, modified seed oil or protein yield and/or |
| CC | content, improved yield by modification of carbohydrate, nitrogen or |
| CC | phosphorus use and/or uptake, by modification of photosynthesis or by |
| CC | providing improved plant growth and development under at least one stress |
| CC | condition, improved lignin production or improved galactonmanan |
| CC | production. This sequence represents a bacterial polynucleotide used in |
| CC | the scope of the invention. Note: The sequence data for this patent did |
| CC | not form part of the printed specification but was obtained in electronic |
| CC | format from USPTO at seqdata.uspto.gov/sequence.html. |
| XX | |
| SQ | Sequence 960 BP, 295 A, 196 C, 278 G, 191 T, 0 U, 0 Other: |
| | |
| | Alignment Scores: |
| | Pred. No.: 1,55e-17 Length: 960 |
| | Score: 262.00 Matches: 90 |
| | Percent Similarity: 47.15% Conservative: 59 |
| | Best Local Similarity: 28.46% Mismatches: 113 |
| | Query Match: 11.47% Indels: 55 |
| | DB: 13 Gaps: 14 |
| | |
| | US-10-701-200-6 (1-437) x ADS48162 (1-960) |
| OY | 33 LysLVsValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAnsSerAlaIleGly 52 |
| DB | 4 AAGAAGATAGACGACTTATTACAAGCGGGGAGACGCACCTCGAATAAACACGACTTGAGA 63 |
| OY | 53 SerLeuIIleGIuARgTYrrThrcGuIleAapProSerIIleGluIleIleCystrYrArgGly 72 |
| DB | 64 GCCGGGTTC---AGGTATC---GGTGTCAAGCAGGAGCTGAAGTAGTCCGAGTGAAGA 117 |
| OY | 73 GIYYTrLYSGLIyLeuLeuLIGLyAsPSeTYrrProValThrAlaGluValArGLySLyS 92 |
| DB | 118 GGTTACTCAGGCCCTCATCGACGGCGATTT-----GTAAACTCGATGACAAAGATGTG 171 |
| OY | 93 AlaGIyValLeuGIuARgPhcGIyGlySeValIIleGIyAsnSerXrGVallySLeuThr 112 |
| DB | 112 GCAGGAATCACAGAAAAG---GAGAGAACAAATTTCAGAACTTCAGATGTGAGAGTTTC 228 |
| OY | 113 AsnValILySApPCyVALySaRGLyLeuValLySGuICIyGlnAsPrOGInLyVal 132 |
| DB | 229 AAGACA-----GAAAGGGCAGGGAATC 252 |
| OY | 133 ALaaIAaSPGInLeuValLySaPGLyValaPIleLeuHISrThrIIleGIyGIyaAPaSP 152 |
| DB | 253 GCTGGCAAACACATAAAAAACAATGTAATAGAAGAAGACTGCTGCTAATAGTGTGAAAGG 312 |
| OY | 153 ThrAGnThrAlaAlaAlaSpLeuAlaAlaPhLeuAlaARGAsnAnTYrrGIyLeuThr 172 |

Db 313 AGCTCAAC-----GGCGCTACTCTTTACGAAGAACACAAATATCCC 357
Qy 173 ValIleGIyLeuProIyThrValIlePheAsnAerPValPheProIleIySgInserLeuGIy 192
Db 358 GTTGCGGTATACCAAGAACCATAGACACAGACATGGGTTGATCGACATGCGATGGG 417
Qy 193 AlATrPThrAlaIaIaGlInGInGlyAlaArGlyrPheMetAerValAlaIaGluaAsn 212
Db 418 GTGACACACGTGTTGAACAACGGGTATGATGCTGTTCAAAAGCTCAAGACACCGTAGC 477
Qy 213 AlaAsnProArGmetLeuIleValIleSgIuValImetGIyArGAsnCySgIyTrPLeuThr 232
Db 478 TCGCATAGAGACCTTTCATTGG--GAAGTCATCGGGAAGGCATTCGGTCATATGCT 534
Qy 233 AlaAlaThrAlaGInGInuIyArGlyrIleLeuAerPAlaGInIyTrPLeuProGluLeu 253
Db 535 CTCATGCGGGGA-----CTGGGAGCTGTGTGCAGAA----- 564
Qy 253 GIyLeuThrArGInuSerTyrGIuValIleAlaValPheAlProGluMetAlaIleAsp 272
Db 565 -----GCCATCATCTGTCACAGAGATTCGCGTGAT 594
Qy 273 LeuGluAlaGluAlaIyArGleuArGInuVal-----MetAspIyValAspCyVal 290
Db 595 TATTCACAGCTGCGCATGAATTCTCGAAGAAAGAGAAGAAAGATATCAACAGATA 654
Qy 291 AsnIlePheValIerGIuGIyAlaGIyAlaIleValAla--GluMetGlnAla 309
Db 655 ---ATCATAGTCCGTAGAGGGGACAGAGTCCCTTACCGTCGCAAGACACCTCGAATAC 711
Qy 310 Lys---GIyGInGInuValProArGAspAlaPheGIyHISileIyIyLeuAspAlaValAsn 328
Db 712 AGGATAGGCTACGAAACGAGATCAACCATATCTCGACACGACAGAG----- 758
Qy 329 ProGlyLysTrPheGIyGIyGInGInPheAlaGmetIleGIyAlaGIu 344
Db 759 -----AGGTGTCTTCACAGCGCTTTCAGCAGAAAGACTCGCATGAG 800

Search completed: March 1, 2005, 15:19:45
Job time : 776 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 1, 2005, 13:55:46 ; Search time 235 Seconds
(without alignments)
3042.778 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284

Sequence: 1 DVTWPHYLRADIRCFHMF.....FNSMLSEIQPKGKGVESH 437

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 segs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DRV=hlh
-Q/cgn2_1/USFTO_spool/US10701200/runat_25022005_102833_10340/app_query.fasta_1.583
-DB=Issued_Patente_NA -OFMT=fasta -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10701200@cgn1_1.1.69@runat_25022005_102833_10340 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEBUERY -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patente NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUTS COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfillseq.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 2284 | 100.0 | 1311 | 4 | US-09-934-901-15 |
| 2 | 2284 | 100.0 | 1311 | 4 | US-09-934-868-5 |
| 3 | 2284 | 100.0 | 1311 | 4 | US-10-321-210-15 |
| 4 | 2284 | 100.0 | 1311 | 4 | US-10-320-874-15 |
| 5 | 270 | 11.8 | 1778 | 1 | US-08-416-870C-3 |
| 6 | 264.5 | 11.6 | 1624 | 1 | US-08-416-870C-5 |
| 7 | 262 | 11.5 | 1978 | 1 | US-08-416-870C-1 |
| 8 | 261 | 11.4 | 1038 | 4 | US-09-902-540-8974 |
| 9 | 261 | 11.4 | 10391 | 4 | US-09-902-540-958 |
| 10 | 260 | 11.4 | 1558 | 1 | US-08-416-870C-9 |
| 11 | 246.5 | 10.8 | 2964 | 4 | US-09-514-221A-404 |
| 12 | 239.5 | 10.5 | 6196 | 4 | US-09-902-540-788 |

| | | | | | | |
|----|-------|------|---------|---|---------------------|--------------------|
| 13 | 228 | 10.0 | 7972 | 4 | US-08-956-171E-312 | Sequence 312, App |
| 14 | 228 | 10.0 | 7972 | 4 | US-08-781-986A-312 | Sequence 312, App |
| 15 | 227 | 9.9 | 1017 | 4 | US-09-489-039A-5087 | Sequence 5087, App |
| 16 | 226.5 | 9.9 | 1038 | 3 | US-09-134-001C-937 | Sequence 937, App |
| 17 | 226.5 | 9.9 | 4244 | 4 | US-09-710-279-4256 | Sequence 4256, App |
| 18 | 223 | 9.8 | 4403765 | 3 | US-09-103-840A-2 | Sequence 2, Appli |
| 19 | 223 | 9.8 | 4411529 | 3 | US-09-103-840A-1 | Sequence 1, Appli |
| 20 | 222.5 | 9.7 | 1008 | 4 | US-09-583-110-2438 | Sequence 2438, App |
| 21 | 222.5 | 9.7 | 1086 | 4 | US-09-902-540-7871 | Sequence 7871, App |
| 22 | 220 | 9.6 | 993 | 4 | US-09-543-681A-4165 | Sequence 4165, App |
| 23 | 218.5 | 9.6 | 19446 | 3 | US-08-961-527-51 | Sequence 1927, App |
| 24 | 217.5 | 9.5 | 19446 | 3 | US-08-961-527-51 | Sequence 51, Appl |
| 25 | 215.5 | 9.4 | 972 | 4 | US-09-107-532A-2816 | Sequence 2816, App |
| 26 | 212.5 | 9.3 | 4605 | 3 | US-09-221-017B-128 | Sequence 128, App |
| 27 | 211.5 | 9.3 | 2048 | 1 | US-08-416-870C-7 | Sequence 7, Appli |
| 28 | 210.5 | 9.2 | 640681 | 4 | US-09-990-988-1 | Sequence 1, Appli |
| 29 | 210 | 9.2 | 2510 | 4 | US-09-949-016-3617 | Sequence 3617, App |
| 30 | 208.5 | 9.1 | 2880 | 4 | US-09-614-221A-354 | Sequence 354, App |
| 31 | 206.5 | 9.0 | 3047 | 1 | US-08-280-690-1 | Sequence 1, Appli |
| 32 | 204.5 | 9.0 | 2591 | 4 | US-09-023-655-1414 | Sequence 1414, App |
| 33 | 204.5 | 9.0 | 2591 | 4 | US-09-814-915A-102 | Sequence 102, App |
| 34 | 204.5 | 9.0 | 2591 | 4 | US-09-949-016-380 | Sequence 380, App |
| 35 | 203.5 | 8.9 | 1274 | 4 | US-09-887-054-1 | Sequence 1, Appli |
| 36 | 202.5 | 8.9 | 1830121 | 4 | US-09-557-884-1 | Sequence 1, Appli |
| 37 | 202.5 | 8.9 | 1830121 | 4 | US-09-643-990A-1 | Sequence 5225, App |
| 38 | 201.5 | 8.8 | 2509 | 4 | US-09-949-016-5225 | Sequence 1, Appli |
| 39 | 195 | 8.5 | 580073 | 4 | US-08-545-528B-1 | Sequence 1, Appli |
| 40 | 182 | 8.0 | 924 | 4 | US-09-710-924-2121 | Sequence 2121, App |
| 41 | 181 | 7.9 | 1230025 | 4 | US-09-198-452A-1 | Sequence 1, Appli |
| 42 | 181 | 7.9 | 1230230 | 4 | US-09-438-185A-1 | Sequence 1, Appli |
| 43 | 179 | 7.8 | 1230230 | 4 | US-09-438-185A-1 | Sequence 1, Appli |
| 44 | 176.5 | 7.7 | 2255 | 3 | US-08-714-918-105 | Sequence 105, App |
| 45 | 176.5 | 7.7 | 2255 | 3 | US-09-265-315-105 | Sequence 105, App |

ALIGNMENTS

RESULT 1
US-09-934-901-15
; Sequence 15, Application US/09934901
; Patent No. 655353
; GENERAL INFORMATION:
; APPLICANT: Koffas, Matheos
; APPLICANT: Odom, J. Martin
; APPLICANT: No. 6555353ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 901
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: METHYLOMONAS SP.
US-09-934-901-15

Alignment Scores:

Pred. No.: 7,48e-264 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-701-200-6 (1-437) x US-09-934-901-15 (1-1311)

QY 1 AagvaValThrTrpProGlyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe 20
DB 1 GATGTGTCATCATGGCCCTATCACTTAACGGCTGATGATTCATTGTGTCATGTTT 60

QY 21 LeuAenPhaAenPhetYrThrleuWetAenlySProlYleYsValAlaileuThraJa 40
DB 61 CTTAACTTAACTTACACAGCTCAAGAACAACTTAAAGATTGCAATACAGACAGA 120
QY 41 GlyValLeuAlaProCySeleuAenSerAlaileGlySerleuileGluYrgrYrThrglu 60
DB 121 GGCGGCTGGCCCTGTTGTAATTCGCAATCGAGTGAATTGATTCGAACGTTATACCGAA 180
QY 61 lAeAPProSerileGluileileCySTYrArgGlyGlyTYrlyeGlyleuLeuengly 80
DB 181 ATCGATCTAGCATGAATCAATTTGCTATCGGGCGGTTATTAAGCCCTGTTGCTGGGC 240
QY 81 AepSerTYrProValThrAlaGluValAlaGlylyblybAlaGlyValleuGluYrphGly 100
DB 241 GATTTCTTATCCAGTAACGCCCAAGCGCTTAAAGCGGCTTCTGCAAGCTTTGGC 300
QY 101 GlySerValileGlyAenSerArgValileYblyeUThraenVallybAPCyVallybA 120
DB 301 GGTTCGTGATCGGCAACGCCGCTCAATTTGACCAATGTCAAAGACTGCGTGAACGC 360
QY 121 GlyLeuVallyeGlyGlyGluYrProGluYrValAlaAlaAPGluYrleuVallybAP 140
DB 361 GGTTCGTGTAAGAGGCTGAAGATCCGCAAAAGTCGGCGCTGATCAATTTGCTTAAGAT 420
QY 141 GlyValAepIleuUHisThrileGlyGlyAepAPerThrAenThrAlaAlaAepIleu 160
DB 421 GGTGTCGATATTCGACACACATCGCGCGGATGATACCAATACGACACAGCGATTTG 480
QY 161 AlaAlaPhleuAlaArgAenAenTYrGlyleuUThraileGlyleuProlySerThrVal 180
DB 481 GCAAGCATCTCGCCGCAAAATTAATTAAGGATGACCGTATGTTTACCTTAACCGTC 540
QY 181 AepAenAPValPhaProIleYsGlnSerleuGlyAlaTrpThrAlaAlaGluGlnGly 200
DB 541 GATTAACGATATTTCCATCAAGCAATCACTAGTGCTTGAATCGCGCGGCAAGGC 600
QY 201 AlaArgTYrPhenecaAenValAlaAlaGluAenAenAlaAenProArgHeleuileVal 220
DB 601 GCGCGTTATTTATCATGAACGTGGTGGCGCAAAACAAGCCCAACCGCATGCTGATCGTA 660
QY 221 HsGluValMetGlyArgAenCySGlyTYrPleuUThraileAlaThrAlaGlnGluTYrArg 240
DB 661 CACGAAGTATAGGCGCGTAATCTGCGCTGCTGACCGCTGCAACCGCGCAAGAAATTCGC 720
QY 241 LybLeuLeuAPArgAlaGluYrPleuProGluYrleuGlyleuThraGlnSerTYrGlu 260
DB 721 AATTAATCTGACCGTGCAGTGGTGGCGGAATGGTGGTGAATCTGTAATCTTAATGA 780
QY 261 ValHsAlaValAlaPheValProGluWetAlaileAepIleuGluAlaGluAlaYbArgleu 280
DB 781 GTGCAAGCGGTATTCCTTCGCAAAATGGCGATCGACTGGAAGCCCAACCGCAAGCCCTCG 840
QY 281 AArgGluValMetAepIleYblyeValAepCyblyeValAenIlePheValSerGluYrAlaGlyVal 300
DB 841 CCGGAAGTATAGCAAAAGTCATGCGTCAACATTTGTTTCGAAGGTCGCGCGCTC 900
QY 301 GluAlaileValAlaGluWetGlnAlaYsGlyGlnGluValProArgAenAlaPheGly 320
DB 901 GAAGCATATGTCGCGAAATGCAAGCGCAAAAGCGCAAGAGTCCGCGCATCGTTCCGC 960
QY 321 HsIleYblyeUThraValAlaAenProGlyTYrPheGlyGlyGluGlnPheAlaGlnMet 340
DB 961 CACATCAAACTGATCGCGTCAACCTCGTAAATGTTGCGGAGCAAAATTCGCGCAGAG 1020
QY 341 lAeGlyAlaGluYrThrleuValAlaGlyYsSerGlyTYrPheAlaArgAlaSerAlaSer 360
DB 1021 ATAGGCGCGAAACAACTCGGTAACAAATCGGCTACTTCCCGCTCTCTGCTTC 1080
QY 361 AenValAepAPerAepIleYblyeUThraAenSerCybAlaAepIleuAlaGluYrAlaPhe 380
DB 1081 AACGTTGACGATGCTGTTGATCAAAATGCTGCGCGCATTTGGCGGTCAAGTCCGCGTTC 1140

QY 381 ArgArgGluSerGlyValileGlyYblybAPGluAepAenGlyAenValleuArgAlaile 400
DB 1141 GGCCCGGAGTCTGGGCTATCGTTCACAGCAAGCAACCGCAAGCTGTTGGTGCAGATC 1200
QY 401 GluPheProArgileYleGlyGlyYbSerProPheAenIleAepThrAepTrpPheAenSer 420
DB 1201 GAGTTCCGCGCATACAGGCGGCAACCGCTTCAATATCGACACCGATCGTTCATTAAC 1260
QY 421 MetLeuSerGluileGlyGlnProlySGlyGlylybValAlaGluValSerHis 437
DB 1261 ATGTTGACGAAATCGCCGACGCTTAAGCGGTTAAAGTCAAGTCAAGCCAC 1311
RESULT 2
US-09-934-868-5
/ Sequence 5, Application US/09934868
/ Patent No. 6689601
/ GENERAL INFORMATION:
/ APPLICANT: Koffas, Mattheos
/ APPLICANT: Odem, James M
/ APPLICANT: Schenzle, Andreas J
/ TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
/ FILE REFERENCE: CL1596 US NA
/ CURRENT APPLICATION NUMBER: US/09/934,868
/ PRIOR APPLICATION NUMBER: 60/229,858
/ PRIOR FILING DATE: 2000-09-01
/ NUMBER OF SEQ ID NOS: 81
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 5
/ LENGTH: 1311
/ TYPE: DNA
/ ORGANISM: METHYLOMONAS SP.
US-09-934-868-5
Alignment Scores:
Pred. No.: 7,486-264 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-10-701-200-6 (1-437) x US-09-934-868-5 (1-1311)
QY 1 AepValAlaThrTrpProTYrHisleuThraAepIleArgPheCyHisTrpPhePhe 20
DB 1 GATGGTCACATGCGCCCTATCACTTAACGGCTGATTCGATTTGTCATGTTT 60
QY 21 LeuAenPhaAenPhetYrThrleuWetAenlySProlYblybValAlaileuThraJa 40
DB 61 CTTAACTTAACTTACACAGCTCAAGAACAACTTAAAGATTGCAATACAGACAGA 120
QY 41 GlyValLeuAlaProCySeleuAenSerAlaileGlySerleuileGluYrgrYrThrglu 60
DB 121 GGCGGCTGGCCCTGTTGTAATTCGCAATCGAGTGAATTGATTCGAACGTTATACCGAA 180
QY 61 lAeAPProSerileGluileileCySTYrArgGlyGlyTYrlyeGlyleuLeuengly 80
DB 181 ATCGATCTAGCATGAATCAATTTGCTATCGGGCGGTTATTAAGCCCTGTTGCTGGGC 240
QY 81 AepSerTYrProValThrAlaGluValAlaGlylyblybAlaGlyValleuGluYrphGly 100
DB 241 GATTTCTTATCCAGTAACGCCCAAGCGCTTAAAGCGGCTTCTGCAAGCTTTGGC 300
QY 101 GlySerValileGlyAenSerArgValileYblyeUThraenVallybAPCyVallybA 120
DB 301 GGTTCGTGATCGGCAACGCCGCTCAATTTGACCAATGTCAAAGACTGCGTGAACGC 360
QY 121 GlyLeuVallyeGlyGlyGluYrProGluYrValAlaAlaAPGluYrleuVallybAP 140
DB 361 GGTTCGTGTAAGAGGCTGAAGATCCGCAAAAGTCGGCGCTGATCAATTTGCTTAAGAT 420
QY 141 GlyValAepIleuUHisThrileGlyGlyAepAPerThrAenThrAlaAlaAepIleu 160

```

Db      421 GGTGTGATATCTGACACCATCGCGGATGAATCCAAATCGGCAGACGGATTGG 480
Qy      161 ALaAlaPheLeuAlaArgAenAenTYrGlyLeuThrValIleGlyLeuProlysthrVal 180
Db      481 GCAGCATTCCTGGCCAGAAATATTAACGACTGACCGCTCATGGTTACTTAACACCGTC 540
Qy      181 AspAenAspValPheProIlelysgInserLeuGlyValATrThrAlaIleGInGInGly 200
Db      541 GATTAACGACGATATTCGATCAAGCAATCACTAGTGCTTGACCTCGCGGAGCAAGGC 600
Qy      201 ALaArgTYrPheMetAenValValAlaGluAenAenAlaAenProArgMetLeuIleVal 220
Db      601 GCGCGTTATTTCAATGACGCGGCGGCGAAACACGCCAACCGCATGCGATCGTA 660
Qy      221 HLeGluValMetGlyYrGAnCySGlyTYrPLeuThrAlaAlaThrAlaGInGInTYrArg 240
Db      661 CACGAAGTATGGGCGCGTAATCGCGCTGCTGACCGCTGCAACCGCGCAGAAATTCGC 720
Qy      241 LysLeuLeuAspArgAlaGluTYrPLeuProGluLeuGlyLeuThrArgGInserTYrGlu 260
Db      721 AATTAATCTGGACCGTCCGAGTGCTTCCGCGAAATGGGTTGACTCGTAATCTTAAGA 780
Qy      261 ValHLeAlaValPheValProGInMetAlaIleAspLeuGluAlaGluAlaIleValArgLeu 280
Db      781 GTGCACGCGGATTCGTTCCGGAATGCGCATGCACTGGAACCGACGCAAGCCGCTCG 840
Qy      281 ArgGluValMetAspArgValAspCysValAenIlePheValserGInGlyAlaGlyVal 300
Db      841 CCGGAAGTATGACAAAGTCATCGATCGTCACATCTTCCTTCGCAAGGTGCGGCGTC 900
Qy      301 GluAlaIleValAlaGluMetGInAlaIleGlyGInGluValProArgAspAlaPheGly 320
Db      901 GAAGCATATCGTCCGGAATATGACGCGCAAGCCAGAAAGTGCAGCGCATGCTTCGGC 960
Qy      321 HLeIleLysLeuAspAlaValAsnProGlyLysTYrPheGlyGluGInPheAlaGInMet 340
Db      961 CACATCAAACTGATCGGTCACCCCTGTGAATGTTCGCGCAGCAATTCGCGCAGATG 1020
Qy      341 ILeGluAlaGluYrThrLeuValGInLysSerGlyTYrPheAlaArgAlaSerAlaSer 360
Db      1021 ATTAGCGCGGAATAAAACCTTGATCAAAATCCGGAATCTTCGCGCTGCTTCCTCC 1080
Qy      361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
Db      1081 AACGTTAGACACATGGCTTTGATCAATGCTGCGCGCATGTGGCGGTCCAGTCCGGCTTC 1140
Qy      381 ArgArgGluSerGlyValIleGlyHLeAspGluAspAsnGlyAenValLeuArgAlaIle 400
Db      1141 CGCCGCGAGCTGCGCTGATCGGTCAACGAAAGCAACCGCAACGTTGCTGCGCATC 1200
Qy      401 GluPheProArgIleLysGlyLysProPheAsnIleAspThrAspTYrPheAsnSer 420
Db      1201 GAGTTTCGCGCATCAAGGGCGCAACGTTAAATATTCACACCGCATGTTCAATAGC 1260
Qy      421 MetLeuSerGluIleGlyGInProLysGlyLysValGluValserHis 437
Db      1261 ATGTTAGCGAAATCGGCCAGCTTAAGCGGTAAAGTCCAACTCAGCCAC 1311

```

```

; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: METHYLOMONAS SP.
US-10-321-210-15

Alignment Scores:
Pred. No.: 7.48e-264 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Gaps: 0
DB: 4

US-10-701-200-6 (1-437) x US-10-321-210-15 (1-1311)
Qy      1 AspValAlaThrTrpProIleHLeuThrAlaAspIleArgPheCysHLeuTrpPhe 20
Db      1 GATGTGTCACATGSCCTTAATCACTTAACGCGCTGATATTCGATTTGTCAATGGTTT 60
Qy      21 LeuAsnPheAsnPheTYrThrLeuMetAenLysProLysLysValAlaIleLeuThrAla 40
Db      61 CTTAACCTTAATCTTACACGCTCAACGCTCAAGAACCAACCTTAAGAAAGTTGCAATACGACGCA 120
Qy      41 GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTYrThrGlu 60
Db      121 GCGGCGCTTGGCGCTTGTGAATTCGCAATCGGATGTTGATCCAACTTAATACGAA 180
Qy      61 IleAspProSerIleGluIleIleCysTYrArgGlyGlyTYrLysGlyLeuLeuLeuGly 80
Db      181 ATCGATCTTAAGCAATAAATCAATTCATTCGAGGCGGTATTAAGCCCTGTTGCTGGGC 240
Qy      81 AspSerTYrProValThrAlaGluValArgLysLysAlaGlyValLeuGInArgPheGly 100
Db      241 GATTCCTTAACGATACGCGGCGCAAGTCCGTAAGACCGGGGTCTTGCACGTTTGGC 300
Qy      101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
Db      301 GGTTCGTATATGCGCAACGCGCGCTCAATTCACCAATGTCAAGACTGCGTGAACGC 360
Qy      121 GlyLeuValLysGluGlyLysAspProGInLysValAlaAlaAspGInLeuValLysAsp 140
Db      361 GGTTCGTAAAGAGGATGAAGATCCGCAAAAGTCCGCTGATCAATTTGTTAAGAT 420
Qy      141 GlyValAspIleLeuHLeuHLeuIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIle 160
Db      421 GGTTCGATATTCGACACCATCGCGCGCGATGATCAATACGCGCAGCGGATTGG 480
Qy      481 GCAGCATTCCTGGCCAGAAATATTAACGACTGACCGCTCATGGTTACTTAACACCGTC 540
Db      161 ALaAlaPheLeuAlaArgAenAenTYrGlyLeuThrValIleGlyLeuProlysthrVal 180
Qy      181 AspAenAspValPheProIlelysgInserLeuGlyValATrThrAlaIleGInGInGly 200
Db      541 GATTAACGACGATATTCGATCAAGCAATCACTAGTGCTTGACCTCGCGGAGCAAGGC 600
Qy      201 ALaArgTYrPheMetAenValValAlaGluAenAenAlaAenProArgMetLeuIleVal 220
Db      601 GCGCGTTATTTCAATGACGCGGCGGCGGAAACACGCCAACCGCATGCGATCGTA 660
Qy      221 HLeGluValMetGlyYrGAnCySGlyTYrPLeuThrAlaAlaThrAlaGInGInTYrArg 240
Db      661 CACGAAGTATGGGCGCGTAATCGCGCTGCTGACCGCTGCAACCGCGCAGGAATTCGC 720
Qy      241 LysLeuLeuAspArgAlaGluTYrPLeuProGluLeuGlyLeuThrArgGInserTYrGlu 260
Db      721 AATTAATCTGACCGTCCGAGTGCTTCCGCGAAATGGGTTGACTCGTAATCTTAAGA 780

```


QY 261 VALHISALAVAPHEVALPROGLUMETALAIIEASPLEUGLUALAGLUALALYBARGLEU 280
 DB 781 GGGCGCGGATATGCTTCCGAAATGGGATCGACCTTGAAAGCCGAAGCCGAAAGCCGCTG 840
 QY 281 ARGGLUVALMETASPLVVALASPCYVALAANIIEPHEVALSERGLUGLYVALAGLYVAL 300
 DB 841 CGCGAAGTATGATGACAAAGTCAATTCGCTTCTGTTCCGAAGTCCGGCGCTC 900
 QY 301 GLUALAILEVALALAGLUMETGLNALALYBGLYGLINGLUALPROARGASPLAAPHGLY 320
 DB 901 GAAGCTATCGTCGCGGAAATGCAAGGCCAAGGCAAGAGTCCGCGCATCGTTCCGCG 960
 QY 321 HISILIEVLEUASPLAVAVAIASNPROGLYVSTPPHEGLYGLINGLPHVALAGLUMET 340
 DB 961 CACATCAAACTGATGAGCGGTCAACCTGTAAATGTTCCGAGCAATTCGCGAGATG 1020
 QY 341 IIEGLYVALAGLULYETHRLEUVALGLNYSERGLYTRPHEALARGALASERLASER 360
 DB 1021 ATAGCGCGGAAAAAACCCTGGTACAAAATCGGATACCTTCCCGCTCTCTGCTTC 1080
 QY 361 AENVALASPMETARGLEUIELYSERCYVALAASPLEUALVALGLUCYBALAPHE 380
 DB 1081 AACGTTGACGACATCGTTGATCAATCGTCGCGCATGCGCGATCGAGTCCGCTTC 1140
 QY 381 ARGATRGUSERGLYVALIEGLYHISAPSGIUAASPGIYASNPVALLEUARGALALIE 400
 DB 1141 CCGCGGAGTCTGGCGGTATCGGTCAAGCAAGACAAAGGCAAGCTGTTGCTCGATC 1200
 QY 401 GIUPHEPROARGILEYBGLYGLYBPROPHASPLAASPTHRASPTHRPHEASINER 420
 DB 1201 GAGTTTCGCGCATCAAGGCGGCAAAACCTTCAATATGACACCGATCGTTCAATAGC 1260
 QY 421 METLEUSERGLIIEGLYGLINPROLYBGLYGLYVVALGLUVALSERHIS 437
 DB 1261 ARGTTGACGAATCGGCCAGCCTAAAGCGGTAAAGTGAAGTCAGCCAC 1311
 DB
 RESULT 4
 US-10-320-874-15
 / Sequence 15, Application US/10320874
 / Patent No. 6773905
 / GENERAL INFORMATION:
 / APPLICANT: Kofas, Matheos
 / APPLICANT: Odom, J. Martin
 / APPLICANT: No. 6773905ton, Kelley C.
 / APPLICANT: Ye, Rick
 / TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
 / FILE REFERENCE: CL1619 US NA
 / CURRENT APPLICATION NUMBER: US/10/320,874
 / CURRENT FILING DATE: 2002-12-17
 / PRIOR APPLICATION NUMBER: US/09/934,901
 / PRIOR FILING DATE: 2001-08-22
 / PRIOR APPLICATION NUMBER: 60/229,906
 / PRIOR FILING DATE: September 1, 2000
 / NUMBER OF SEQ ID NOS: 20
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 15
 / LENGTH: 1311
 / TYPE: DNA
 / ORGANISM: METHYLOMONAS SP.
 US-10-874-15
 Alignment Scores:
 Pred. No.: 7,48e-264 Length: 1311
 Score: 2284.00 Matches: 437
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0
 US-10-701-200-6 (1-437) x US-10-320-874-15 (1-1311)
 QY 1 AENVALVALTHRTTPROFYRHISLEUTHRALAASPLIARGPHECYHISERTPPHEPHE 20
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 GATGTCACATGGCCCTATCACTTAACGGCTGATATTCGATTTTGTGATGTTTTT 60
 QY 21 LEUENPHEANPHEYTRTHLEUMECANLYBPROLYBLYVALAILEUTHRALA 40
 DB 61 CTTAACTTTAACTTTCTACACCTCTCATGAACAACTTAATAAGTTGCAATCTGACAGA 120
 QY 41 GLYGLILEUALAPROCYALEUANSERAIIEGLYSERLEUIIEGLIARGYTRHGRU 60
 DB 121 GCGGCTTGCGCTTGTTGTAATTCGCAATCGGATGATGATGACGTTATACCGNA 180
 QY 61 IIEASPPROSERILEGLUIEILECYERTYRARGIYGLYTRLYBGLYLEUENLEUGLY 80
 DB 181 ATGCATCTTACATATGAATCATTTGCTATCGCGCGGTATTAAGGCTGTTGCTGGGC 240
 QY 81 ASPTERTYRPROVALTHRALGLUVALARGLYBLYBALAGLYVALLEUGLARGPHEGLY 100
 DB 241 GATTTTATCCAGTAAACCGGAAAGTGGTAAAGGCGGGGTCTCTCAACGTTTGGC 300
 QY 101 GLYSERVALIIEGLYASNSERARGVALLYSELEUTHRANVALLYASPCYVALLYBARG 120
 DB 301 GGTTCGTGATCGGCACAGCCGCGGTCAAAATTGACCAATGTCAAAGACTGCGTGAACGC 360
 QY 121 GLYLEUVALLYBGLUGLYGLUASPPROGLINLYBVALAIAIASPGLNLEUVALLYBASP 140
 DB 361 GGTTCGTGATCGGCACAGCCGCGGTCAAAATTGACCAATGTCAAAGACTGCGTGAACGC 420
 QY 141 GLYVALASPLIIELEUHIETHRTIEGLYGLYASPSPTHRANTRHALAIAIASPLEU 160
 DB 421 GGTGTCGATATTTGACACACCATCGCGCGCATGATCAATATACCGCGACGCGGATTTG 480
 QY 161 ALAIAIAPHEUVALAARGASPSNTYRGLYLEUTHRVALIIEGLYLEUPROLYBETHRVAL 180
 DB 481 GCGAGATTCCTGGCGCAAGAAATTAATTCGAGCTGACCGCATTTGTTACTTAAACCGTC 540
 QY 181 ASPANASPVVALPHEPROIIEYBGLINSELEUGLYALATRPTHRALAIAAGLUGLNGLY 200
 DB 541 GATAACGACGTRATTTCCGATCAAGCAATCACTAGGTGCTTGACCTGCCGCGAGCAAGGC 600
 QY 201 ALAARGTYRPHMETASNPVALAIAAGLUBENASNAIASNPARGMECLEUIIEVAL 220
 DB 601 GCGCGTTATTTTCAIAGAAAGTGTGCGCCGAAACCAACGCCACGATGCTGATCTGTA 660
 QY 221 HISGLIUALMETGLYARGASNPVCSGLYTRPLEUTHRALAIAETHRALAGLULYTRARG 240
 DB 661 CACGAAAGTATGGGCGTACTGCGGCTGCTGACCGCTGCAACCGCGCAGAAATATGCG 720
 QY 241 LYSLEUENASPARVALAGLUTRPLEUPROGLIUEUGLYLEUTHRARGUSERYRGIU 260
 DB 721 AATTTACTGACCGCGCGGATGGGTGCGGAATTGGTTTGACTCGTGAATCTTATGAA 780
 QY 261 VALHISALAVAPHEVALPROGLUMETALAIIEASPLEUGLUALAGLUALALYBARGLEU 280
 DB 781 GTGCACGCGGTATTCGTTCCGAAATGCGATCGACCTCGGAAGCCGAAGCCGAGCGCTG 840
 QY 281 ARGGLUVALMETASPLVVALASPCYVALAANIIEPHEVALSERGLUGLYVALAGLYVAL 300
 DB 841 CGCGAAGTATGATGACAAAGTCAATTCGCTTCTGTTCCGAAGTCCGGCGCTC 900
 QY 301 GLUALAILEVALALAGLUMETGLNALALYBGLYGLINGLUALPROARGASPLAAPHGLY 320
 DB 901 GAAGCTATCGTCGCGGAAATGCAAGGCCAAGGCAAGAGTCCGCGCATCGTTCCGCG 960
 QY 321 HISILIEVLEUASPLAVAVAIASNPROGLYVSTPPHEGLYGLINGLPHVALAGLUMET 340
 DB 961 CACATCAAACTGATGAGCGGTCAACCTGTAAATGTTCCGAGCAATTCGCGAGATG 1020
 QY 341 IIEGLYVALAGLULYETHRLEUVALGLNYSERGLYTRPHEALARGALASERLASER 360
 DB 1021 ATAGCGCGGAAAAAACCCTGGTACAAAATCGGATACCTTCCCGCTCTCTGCTTC 1080
 QY 361 AENVALASPMETARGLEUIELYSERCYVALAASPLEUALVALGLUCYBALAPHE 380
 DB 1081 AACGTTGACGACATCGTTGATCAATCGTCGCGCATGCGCGATCGAGTCCGCTTC 1140

Db 1231 GGGGGCTACACGGGCTTACAAAGTGGCTTGTCAATGTAGA---CAGACTTATATCCA 1287
 Qy 402 PheProArgIleuArgIleuProPheAsnIleAspThrAsp-----TrpPheAsn 419
 Db 1288 TTTATCGTATCATCTGAGAGCAGATACGTTGTATACCGATGATGGGCAAGG 1347
 Qy 420 SerMetLeuSerGluIleGlyIlePro 428
 Db 1348 ---CTTGTGTATCCACCAACCAACCA 1371

RESULT 6

US-08-416-870C-5
 / Sequence 5, Application US/08416870C
 / Patent No. 5824862

/ GENERAL INFORMATION:
 / APPLICANT: HIYOSHI, TORU
 / APPLICANT: MINE, TOSHIKI
 / APPLICANT: KASAKA, KEISUKE
 / APPLICANT: TYSON, ROBERT HEN
 / APPLICANT: PAGE, ANTHON MILES JOHN
 / TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 / TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTO
 / TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDE
 / NUMBER OF SEQUENCES: 23
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 / STREET: PO BOX 747
 / CITY: FALL CHURCH
 / STATE: VA
 / COUNTRY: USA
 / ZIP: 22040-0747

COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/416,870C
 / FILING DATE:
 / CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 / NAME: MURPHY JR, GERALD M
 / REGISTRATION NUMBER: 28,977
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (703) 205-8000
 / TELEFAX: (703) 205-8050
 / INFORMATION FOR SEQ ID NO: 5:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1624 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: double
 / TOPOLOGY: linear
 / MOLECULE TYPE: cDNA to mRNA
 / FEATURE:

NAME/KEY: CDS
 / LOCATION: 3..1409
 / US-08-416-870C-5

Alignment Scores:
 / Pred. No.: 2.09e-21 Length: 1624
 / Score: 264.50 Matches: 109
 / Percent Similarity: 40.54% Conservative: 56
 / Best Local Similarity: 26.78% Mismatches: 173
 / Query Match: 11.58% Indels: 69
 / DB: 1 Gaps: 16

US-10-701-200-6 (1-437) x US-08-416-870C-5 (1-1624)

Qy 37 ILeuThrAlaGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGlu 56
 Db 222 ATTGTCACTGTGAGAGACTGTGCTGCCTGCACTGACACTGTCTATTAGGAAATTGTTGT 281

Qy 57 ArgTrpThrGluIleAspProSerIleGluIleIleCysTrpArgGlyGlyTrpArgGly 76
 Db 282 GAGCTAAATGACATGATATGCTGATGAGGAGTACTGGAATTCAGGGGTGGATATGAGCT 341
 Qy 77 LeuLeuLeuGlyAspSerTrpProValThrAlaGluValArgLysLysAlaGlyValLeu 96
 Db 342 TTCTATGCTTGTAAACACATTGACTGAGTCCA-----AACAGGTAAACAGACAT 392
 Qy 97 GluArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAsp 116
 Db 393 CACAAAAGGGGTGGAACTGTTCTGGGACATCACGT----- 428
 Qy 117 CysValLysArgGlyLeuValLysGluIleGluAspProGluLysValAlaAlaAspGlu 136
 Db 429 -----GGAGCCCATGACACCATGACATG---CTTACACAC 461
 Qy 137 LeuValLysAspGlyValAlaAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAla 156
 Db 462 ATCCAGATCGGTGATTAATCATCAGTTTATGTAATTTGGTGTGATGTAATCAAGGGGT 521
 Qy 157 AlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTrpGlyLeuThrValIleGlyLeu 176
 Db 522 GCAGAGTGAATTTTGAAGATTTAGAGACGTGGTCAAGGTTCTGAGTCTGGCAAT 581
 Qy 177 ProLysThrValAspAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAla 196
 Db 582 CCNAGAGCATTTGAATGAATGATACAGTAATGACATGCTTTGGTTTGCACACTGA 641
 Qy 197 AlaGluGlnGlyAlaArgTrpPheMetAsnValAlaGluAsnAsnAlaAsnProArg 216
 Db 642 GTTGAAGAGGCCCAACGTGCAATTAATGCTGCTCATGTGAAGCTGAAGCCCGAGAT 701
 Qy 217 MetLeuIleValHisGluValMetGlyValArgAsnCysGlyTrpLeuThrAlaAlaThrAla 236
 Db 702 GGTATAGGCTCTGTAACCTAATGAGTACAGACAGTGGTTTATTGCA----- 749
 Qy 237 GlnGluTrpArgLysLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArg 256
 Db 750 ---CACTATGCTACTTACCACGACA----- 773
 Qy 257 GluSerTrpGluValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGlu 276
 Db 774 -----GACGTGATGTTGTTGTTGATTCAGAGTCACTTCTATCTGGAAGTGA 824
 Qy 277 Ala-----LysArgLeuArgGluValMetAspLysValAspCys 289
 Db 825 GGTGCTTTTATGATTTTGAAGAAGCTGCAAGAGAAATGTCATATGTT----- 878
 Qy 290 ValAsnIlePheValSerGluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAla 309
 Db 879 -----ATCGTTGTTGCCGAGGTCACAGGCGCAAGAACTTATTATGAACAAGAGATCA 932
 Qy 310 LysGlyGlnGluValProArgAspAlaPheGlyHis-----IleLysLeuAspAlaValAsn 328
 Db 933 ATGGGG-----AAAGATCTTCAGCAATTCGATTTCTTGTATGTT----- 974
 Qy 329 ProGlyLysTrpPheGlyGlnGlnPheAlaGlnMetIleGlyAlaGluLysThrLeuVal 348
 Db 975 ---GGTCTTGTGTTATCTCAAAAGATAAAGAGCACTTCAAGAAATCAAGACTATATA 1031
 Qy 349 Gln-----LysSerGlyTrpPheAlaAlaGluAlaSerAlaSerAsnValAspAsp 364
 Db 1032 AATTCAGATATATATGATCTTACATATGATATGATATGATCTTCATTCAGATATCATCTGAC 1091
 Qy 365 MetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArgLysSer 384
 Db 1092 AATGTGATTCACACTGTGGCACACAGGCTGTTTCATGAGCCATGCGCTGATCACT 1151
 Qy 385 Gly-----ValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIleGluPhePro 403
 Db 1152 GGTTCACGCTGGG-----CAAGTAATGTCGGCATTCG---TATATCCGTTTAC 1202
 Qy 404 ArgIle-----LysGlyGlyLysPheProPheAsnIleAspThrAspTrpPheAsnSerMet 421

```

Db 1203 AGGATTCACAGAGAAAGCAGAAACAAAGTTTCAATTACTATGATAGATGTGGCAGAA---CTT 1259
QY 422 LeuSerGluIleGlyInPro 428
Db 1260 CTCTCTCAACCAACCAACCA 1280

RESULT 7
US-08-416-870C-1
: Sequence 1, Application US/08416870C
: Patent No. 5824862
: GENERAL INFORMATION:
: APPLICANT: HIYOSHI, TORU
: APPLICANT: MINE, TOSHIKI
: APPLICANT: KASAKA, KEISUKE
: APPLICANT: TYSON, ROBERT HUM
: APPLICANT: PAGE, ANTHONY MILES JOHN
: TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
: TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDER
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
: STREET: PO BOX 747
: CITY: PALM CHURCH
: STATE: VA
: COUNTRY: USA
: ZIP: 22040-0747
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,870C
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: MURPHY JR, GERALD M
: REGISTRATION NUMBER: 28,977
: REFERENCE/DOCKET NUMBER: 760-195P (PCT)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1978 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 133..1587
: US-08-416-870C-1

Alignment Scores:
Pred. No.: 5.7e-21 Length: 1978
Score: 262.00 Matches: 113
Percent Similarity: 38.41% Conservative: 56
Best Local Similarity: 25.68% Mismatch: 179
Query Match: 11.47% Indels: 92
DB: 1 Gaps: 17

US-10-701-200-6 (1-437) x US-08-416-870C-1 (1-1978)
QY 7 TyrHisLeuThrAlaAspIleArgPheCysHisIleTrpPhePheLeuAsnPheAsnPheTyr 26
Db 388 TATTTCACTGGATGGATGATTCGTGTCCTTGT
417
QY 27 ThrIleMetAlaenlybProlybysValAlaIleLeuThrAlaGlyIleuAlaProCys 46
Db 418 -----ATTGTAACCTGTGGTGGTGTGGCCCTGGG 447

```

47 LeuAenSerAlaIleGIysLeuIleGIuAArgTyThrGluIleAerProSerIleGIu 66
448 CTTAAACACAGGATGACAGAGATGTGTACATAGCTCGATATTATATAGTGGATGCACAA 507
67 IleIleCybTyArxGIyGIyTyTyAGIyLeuIleuIleuGIyAAsPserTyProValThr 86
508 GTCTTTGGATCGATGAGGGCTACAGGGGTTTCTATTCCAAAGATATCATCATTTGACA 567
87 AlaGIuValAArgLyValAGIyValIleGIuAArgPheGIyGIySerValIleGIyAAsn 106
568 CCA-----AAGCTGTTAATGACATTCATTAAGCTGTGTGTAACAATCTTGATGCA 618
107 SerArgValLyLeuThrAenValLyAAsPcyValLyAArgLyLeuValLyGIuGIy 126
619 TCACGA-----GGAGGC 630
127 GIuAerProGIuAAsValAlaAlaAAsPcIleuValLyAAsArgLyValAAsPcIleuHnIv 146
631 CATATACCAACAAAGATT--GTTGACAGCATACAGACCGTGAAATTAATTCAGGTATAT 687
147 ThrIleGIyGIyAAsPArPThrAenThrAlaAlaAAsPcIleuAlaAAsPcIleuAlaAArg 166
688 ATATCCGTGTGTATGAACTCAGAAAGAGACGCTTATATATGAGAAATGACGGCG 747
167 AAsAenTyArgLyLeuThrValIleGIyLeuProLyThrValAAsPAAAsPValPhePro 186
748 CGTGCTCAAGATTAATGTTGCTGGAGATCCAAAGACAATTGATTAATGATATCCGTGT 807
187 IleLyGIuSerLeuGIyAlaATPrPThrAlaAGIuGIuGIyAlaAArgTyPheMetAAsn 206
808 ATCAACAAGCATTTGTGTGATATACGCTGTACAGAGAGGCTCAACGTCATTAATGCA 867
207 ValValAlaGIuAAsnAlaAAsnProAArgMetLeuIleValIleGIuValMetGIyAArg 226
868 GCTCATGTGTGAAGCTGAAGTGCAGAAATGCTATGTGTGTGTCGAAGCTATATGGAGCC 927
227 AAsCyGIyTyTrPLeuThrAlaAlaAThrAlaAGIuTyTrArgLyLeuLeuAAsPArgAla 246
928 TATAGTGATTCATCGCA----- 945
247 GIuTyTrPLeuProGIuLeuGIyLeuThrAArgGIuSerTyGIuValIleValAlaPheVal 266
946 -----ATGTATGCCACTTTGGCGGACAGAGATGTGATCTCTGTTAATT 990
267 ProGIuMetAlaIleAAsPLeuGIuAlaGIuAla-----LyAAsArg 279
991 CCAAGATCACCCCTTTATCTTGAAGAGATGTGTGACTCTTTGAATATCATTTGAAGAAAAG 1050
280 LeuAArgLyValMetAAsPLeuValAAsPcyValAAsnIlePheValSerGIyValAGIy 299
1051 CTCAAAGAAATATGGGCACATGTT-----ATTGATGATGCCGAAGAGCGAGCG 1098
300 ValGIuAlaIleValAlaAGIu--MetGIuAlaAAsGIyGIuGIuValProAArgAAsPAla 318
1099 CAAGAACCTTCTTGCAGAGAGAAATGCCATGCCAAAAACGAACA-----GATGCT 1149
319 PheGIyHisIleLyLeuAAsPAlaValAAsnProGIyTyTrPheGIyGIuGIuAla 338
1150 TCGGGGAACAAGCTTCTCCAGAGATT-----GTTTGTGATTTCCCAAAAAATCAGG 1203
339 GIuMetIleGIyAlaGIuTyThrLeuValGIu-----LySerGIy 352
1204 GATCATTTT--GCTACAAAATAAGATGATGCCATTTACTTTAAGTATATGATCGACAT 1260
353 TyTrPheAlaAAsPAlaSerAlaSerAAsnValAAsPAsPcAArgLeuIleLySerCyAla 372
1261 TACATGATCTGTGTCTTCCAAAGTAATGCCCTGTATAT-----GTATATTGCACT 1311
373 AAsPLeuAlaValGIuCyValAAsPArgAArgGIuSerGIyValIleGIyHisAAsPcIuAAsP 392
1312 CTCTGTCTCAAAAGTGTGTTCATGAGACAAATGCCAGGCTACAGAGGTTTCACTCACTCA 1371

| | | | |
|--|----------|---|------|
| Oy | 393 | -----AanglyAanValleuArgla1leGuPheProAaTg1lelyGlylyAsp | 410 |
| | | | ::: |
| Db | 1372 | CTTGTCATGTGTCGCG---CAGACTATATACCATTCATTCGATTATACCGAACAATAAT | 1428 |
| Oy | 411 | PheAnu1eAspThrAsp-----TrpPheAnSerMetLeuSerGu1leGlyInPro | 428 |
| | | ::: | |
| Db | 1429 | ATGGTGTTATATAACTACAGACAGATGGGACAGT---CTTCCTTGCTCAACCAATCAGCCA | 1485 |
| RESULT 8 | | | |
| US-09-902-540-8974 | | | |
| ; Sequence 8974, Application US/09902540 | | | |
| ; Patent No. 6833447 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Goldman, Barry S. | | | |
| ; APPLICANT: Hinkle, Gregory J. | | | |
| ; APPLICANT: Slater, Steven C. | | | |
| ; APPLICANT: Wisgand, Roger C. | | | |
| ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof | | | |
| ; FILE REFERENCE: 38-10(1)58491B | | | |
| ; CURRENT APPLICATION NUMBER: US/09/902,540 | | | |
| ; PRIOR FILING DATE: 2001-07-10 | | | |
| ; PRIOR APPLICATION NUMBER: 60/217,883 | | | |
| ; PRIOR FILING DATE: 2000-07-10 | | | |
| ; NUMBER OF SEQ ID NOS: 16825 | | | |
| ; SEQ ID NO 8974 | | | |
| ; LENGTH: 1038 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Myxococcus xanthus | | | |
| US-09-902-540-8974 | | | |
| Alignment Scores: | | | |
| Pred. No.: | 2,726-21 | Length: | 1038 |
| Score: | 261.00 | Matches: | 96 |
| Percent Similarity: | 43.90% | Conservative: | 66 |
| Best Local Similarity: | 26.02% | Mismatches: | 141 |
| Query Match: | 11.43% | Indels: | 66 |
| DB: | 4 | Gaps: | 16 |
| US-10-701-200-6 (1-437) x US-09-902-540-8974 (1-1038) | | | |
| Oy | 34 | LyvAla1Ala1leuThra1Ag1yGlyleuA1ProCyLeuAnSerAla1leGlySer | 53 |
| | | | |
| Db | 4 | AAAGTCGCGGTCTACACCGCGGGGGTACTGCCCCGGCTGAACGCGTATCCGCC | 63 |
| Oy | 54 | Leu1leG1uaTgTyrThc1u1leAspProSer1leGu1le1leCyTyrArg1yGly | 73 |
| | | | |
| Db | 64 | GTCGTCGCGCGCCGACACGCC-----CACGGCTTCGAGATGATGGCTCCGAGATG | 117 |
| Oy | 74 | Tyr1yG1yLeu1leu1leuGlyAspSerTyrProValThra1Ag1uValArg1yLyAla | 93 |
| | | | |
| Db | 118 | TGGAAGGGGTGTGTGGAGACACACATCTCCGCTCAGCGGTAAAC-----ACGTCC | 171 |
| Oy | 94 | G1yVal1leuG1naTgPheG1yGlySerVal1leG1yAnSerArgVal-----Lys | 110 |
| | | | |
| Db | 172 | GGAATCTCCACACCGG---GGCGAACAATCTCGGCACCTCGCGGTCAACCGGTCAAG | 228 |
| Oy | 111 | LeuThraAnVal1yAspCyVal1lyArgG1yLeuVal1yG1uG1y1leuAspProG1n | 130 |
| | | | |
| Db | 229 | GTCGAAAACGGGCTGAGACGCGCTCAAGGCGCC----- | 261 |
| Oy | 131 | LyvAla1Ala1leuBGL1euVal1yAspG1yAla1leuN1eThr1leG1yGly | 150 |
| | | | |
| Db | 262 | -----ATCGAACGCAACGGCATCCACGCGGTATGCGCATTTGTGGC | 303 |
| Oy | 151 | AspAspThraAnThra1Ala1AlaAsp1euAla1Asp1euAla1Asp1euAla1Asp1eu | 170 |
| | | | |
| Db | 304 | GAGGCAACGCTGTGCGCGCCACCGCGCATG-----TCGAGAGAGGA | 345 |
| Oy | 171 | LeuThra1leG1yLeuPro1yThrVal1AspAnAspValPhePro1leLyG1nSer | 190 |
| | | | |
| Db | 346 | CTGGGCATCTGCTGCTGCTCGAAGACATCTGCACACATCAACGACACGCACTTCAAC | 405 |
| Oy | 191 | LeuG1yAla1TrpThra1Ala1leG1uG1nG1yAla1ArgTyrPhe-----MetAsnAlaVal | 208 |

| | | | | | |
|--|---|--|-----|---|------|
| D | b | | 466 | TTCGGCTTGCACACCGCCGCTGCCATTCGCCACGAGCATTGACCGGCTGCACCTCCACC | 4655 |
| O | y | | 209 | AlAgLuAenAmEnAlAAAnPrArGrMeUeUlIeValInIGluValMetGlYAdAsnCyS | 228 |
| D | b | | 466 | GCGGAGTGCGCAC-----AACGCGGTCAATCGTCTGGAGGTGTATGGCCCTTCACGTG | 516 |
| O | y | | 229 | GIYTTPLeuThrAlaLaThrAlaGlnGUTrArGLyLeuLeuAAsPArGrAlaGUTrP | 248 |
| D | b | | 517 | GGCTTGATTTGCCACC-----TACGGCGGATCGCTGGCGCGCGCAC--- | 558 |
| O | y | | 249 | LeuProGluLeuGlyLeuThrArgLusertyrgLiValInIAeAlaValAPheValProGU | 268 |
| D | b | | 559 | -----GTATCTGGTGGCGGAG | 576 |
| O | y | | 269 | MetalAlaEprLeuGluAlaGluAlaValaGleu-----ArgLiValMetAAsPlys | 286 |
| D | b | | 577 | ATTCGCCGCACTCCGCGAAGGTGCGCGACATTCACGCGCGCCACGCGGCGGAGCG | 638 |
| O | y | | 287 | ValAAsCySVAlAAmLIePheValSerGluGlyAlaGlyVAlGluAlaIeValAlaGU | 306 |
| D | b | | 637 | ACCTTCTCATT---GTGCTGTGGCGGAGGTATCGCGCATCAAGCTGTCCGCGACAG | 653 |
| O | y | | 307 | MetGlnAlaLyBgLYngLnGuValProArGrAPalaAPheGlyNHIwIeLysLeuAAsPa | 326 |
| D | b | | 694 | CAGAGACAGCTCGTACACGAGCGCGCACATCGACGAGGACGAGCGCGCGCTCGTGGC | 753 |
| O | y | | 327 | ValAAsPrOGlyLySTrPheGlyGluGlnPheAlaGlnMeLIeGlyAlaGUlySThr | 346 |
| D | b | | 754 | GTG-----GGCACCATCTCTGGCGGACGAATAATCGAGCGCGGACCGGCTTCGAG--ACG | 804 |
| O | y | | 347 | LeuValGlnLySerGlyTYTrPheAlaArgAlaSerAlaSerAAsValAAsPArMecArg | 366 |
| D | b | | 805 | CGCGTCTCGCTGCTGGCGGCACATCCAAACGCGCGCGCGCCCCACCGCGCATACCGCGTG | 864 |
| O | y | | 367 | Leu-----LIeLySerCYsAlaAAsPLeuAlaValGluCYsAlaPheArg | 381 |
| D | b | | 865 | CTCGGCACCGCGTACGAGGTCCACCGCTGC--GACATGTGGCGCGCGGAGTTTCGG | 921 |
| O | y | | 382 | ArgGluSerGlyValIIeGlyNHIAAsP | 390 |
| D | b | | 922 | AAGATGGCGCGCTGCGAGCAACGAC | 948 |
| RESULT 9 | | | | | |
| US-09-902-540-958/c | | | | | |
| ; Sequence 958, Application US/09902540 | | | | | |
| ; Patent No. 6833447 | | | | | |
| ; GENERAL INFORMATION: | | | | | |
| ; APPLICANT: Goldman, Barry S. | | | | | |
| ; APPLICANT: Hinkle, Gregory J. | | | | | |
| ; APPLICANT: Slater, Steven C. | | | | | |
| ; APPLICANT: Wiegand, Roger C. | | | | | |
| ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof | | | | | |
| ; FILE REFERENCE: 38-10(1)5849)B | | | | | |
| ; CURRENT APPLICATION NUMBER: US/09/902,540 | | | | | |
| ; PRIOR FILING DATE: 2001-07-10 | | | | | |
| ; PRIOR APPLICATION NUMBER: 60/217,883 | | | | | |
| ; PRIOR FILING DATE: 2000-07-10 | | | | | |
| ; NUMBER OF SEQ ID NOS: 16825 | | | | | |
| ; SEQ ID NO 958 | | | | | |
| ; LENGTH: 10391 | | | | | |
| ; TYPE: DNA | | | | | |
| ; ORGANISM: Myxococcus xanthus | | | | | |
| US-09-902-540-958 | | | | | |
| Alignment Scores: | | | | | |
| Pred. No.: 1,036-19 length: 10391 | | | | | |
| Score: 261.00 Matches: 96 | | | | | |
| Percent Similarity: 43.90% Conservative: 66 | | | | | |
| Best Local Similarity: 26.02% Mismatches: 141 | | | | | |
| Query Match: 11.43% Indels: 66 | | | | | |
| DB: 4 Gaps: 16 | | | | | |

US-10-701-200-6 (1-437) x US-09-902-540-958 (1-10391)

QY 34 LysValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53
 DB 3602 AAAGTGGCCGCTGTCACCGGCGGGGTGACTGCCCGGCTGAACCCGCTCATCCGCGGC 3543
 QY 54 LeuIleGluArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgIleGly 73
 DB 3542 GTTCGTCGCGCGCCCAACGCC-----CACGCTTGAGATGATGCGCTCCGAGATGAT 3489
 QY 74 TyrLysGlyLeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLysLysAla 93
 DB 3488 TGGAGGGGTTGTTGGAGAGACAACCTTCGCTCACCGCTGAACCC-----ACGTC 3435
 QY 94 GlyValIleuGlnArgPheGlyGlySerValIleGlyAsnSerArgVal-----Lys 110
 DB 3434 GGAATCTCCACCGG---GGCGGAACCATCTCGGCACTCGCGGCTCAACCCGTTCAAG 3378
 QY 111 LeuThrAsnValIleAspCysValIleValArgIleGlyLeuValLysGluGlyLeuAspProGln 130
 DB 3377 GTCGAAMACGGGCTGGAGCGCGCTCAAGCGCGCC----- 3345
 QY 131 LysValAlaAlaAspGlnLeuValLysAspGlyValAspIleuValThrIleGlyGly 150
 DB 3344 -----ATCGAACCAACGGCATCCAGCGCTCACCGCTCATCCGCTGATGATGAT 3303
 QY 151 AspAspThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrGly 170
 DB 3302 GAAGGACGCTGTCGCGCGCCGCGCCGCGCATG-----TCGCAAGAGAGA 3261
 QY 171 LeuThrValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSer 190
 DB 3260 CTGCGCATGCTGGGTGTCGCGAAGACCATCGAACAGACATCAACCGCAGCTTCAC 3201
 QY 191 LeuGlyAlaIleThrAlaAlaGluGlnGlyAlaArgTyrPhe-----MetAsnVal 208
 DB 3200 TTCCGCTTGCACCGCGCTCCGCTCCGATCCGACCGAGCCATTCGACCTCCACC 3141
 QY 209 AlaGluAsnAsnAlaAsnProArgMetLeuIleValIleGlyValMetGlyArgAsnCys 228
 DB 3140 GGGGACTCCAC-----AAGCCGCTCATCTGTCGAGGTGATGGGCGCTCACGTC 3090
 QY 228 GlyTyrPLeuThrAlaAlaThrAlaGlnGlyTyrArgLysLeuLeuAspArgIleGluTyr 248
 DB 3089 GGCTGATTCGACAC-----TACGCGGCGATCGCTGCGCGCGGAC----- 3048
 QY 248 LeuProGluLeuGlyLeuThrArgGluSerTyrGluValIleAlaValPheValProGlu 268
 DB 3047 -----GTCTCTCTGGTGGCGGAG 3030
 QY 269 MetAlaIleAspLeuGluAlaGluAlaLysArgLeu-----ArgGluValMetAspLys 286
 DB 3028 ATTCGCGCGGACCTCGCGAAGTGGCGAGACATCAAGCGCGCCGACCGCGCGGCGCC 2970
 QY 287 ValAspCysValAsnIlePheValSerGlyGlyAlaGlyValGluAlaIleValAlaGlu 306
 DB 2969 ACCTTCTCATTT---GTGCTGTGGCGGAGGATACCGCATCAAGCTGTGCGGCGACAG 2913
 QY 307 MetGlnAlaLysGlyGlnGluValProArgAspAlaPheGlyHisIleLysLeuAspAla 326
 DB 2912 CAGGAGCAGCTGTCAACGCGCGCATCTGAGAGAGAGCGCGCGCGCTCGGTGTC 2853
 QY 327 ValAsnProGlyLysThrPheGlyGluGlnPheAlaGluMetIleGlyAlaGluLysThr 346
 DB 2852 GTG-----GGCACCATCTCGGCGCAGAAATGAGCGCGCGCACCGGCTTCGAG---ACG 2802
 QY 347 LeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValAspPheMetArg 366
 DB 2801 CCGGTGTCTGTGGCGACATCAACGCGCGCGCGCGCGCGCGCATGACCGCGTC 2742
 QY 367 Leu-----IleLysSerCysAlaAspLeuAlaValGluCysAlaPheArg 381
 DB 2741 CTCGCCACCGCTACGCGCTGC---GACATGATGGCGCGCGCGGATTCGCG 2685

QY 382 ArgGluSerGlyValIleGlyHisAsp 390
 DB 2684 AAGATGGCCGCTGCGAGAGCAACGAC 2658

RESULT 10
 US-08-416-870C-9
 : Sequence 9, Application US/08416870C
 : Patent No. 5824862
 : GENERAL INFORMATION:
 : APPLICANT: HIYOSHI, TORU
 : APPLICANT: KASAKURA, KEISUKE
 : APPLICANT: TYSON, ROBERT HUM
 : APPLICANT: PAGE, ANTHONY MILES JOHN
 : TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
 : TITLE OF INVENTION: 1-PHOSPHORANSEASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
 : TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDER
 : NUMBER OF SEQUENCES: 23
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
 : STREET: PO BOX 747
 : CITY: FALL CHURCH
 : STATE: VA
 : COUNTRY: USA
 : ZIP: 22040-0747
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : OPERATING SYSTEM: IBM PC compatible
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/416,870C
 : FILING DATE:
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: MURPHY JR, GERALD M
 : REGISTRATION NUMBER: 28,977
 : REFERENCE/DOCKET NUMBER: 760-195P (PCT)
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (703) 205-8000
 : TELEFAX: (703) 205-8050
 : INFORMATION FOR SEQ ID NO: 9:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1558 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: CDNA to mRNA
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 1..1278
 : US-08-416-870C-9

Alignment Scores:
 Pred. No.: 6.8e-21 Length: 1558
 Score: 260.00 Matches: 107
 Percent Similarity: 41.15% Conservative: 65
 Best Local Similarity: 25.60% Mismatches: 174
 Query Match: 11.38% Indels: 72
 DB: 1 Gaps: 16

US-10-701-200-6 (1-437) x US-08-416-870C-9 (1-1558)

QY 37 IleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGlu 56
 DB 148 ATTGTACATGTGTGGCTTGTGTCACAGGCTTAATACATGATCAGAGAAATCGTTGT 207
 QY 57 ArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGly 76
 DB 208 GGAATGTCTTACATGATGATGCTCAAGAAATCTTGGCATTCAGAGGAGCTTACAGAGC 267
 QY 77 LeuLeuLeuGlyAspSerTyrProValThrAlaGluValArgLysValAlaGlyValLeu 96

[illegible]

```

Db      1186  ---CTTTGCTTCGACAAACAGCCCTAGTTCTATGAAGACGATGATCACACC 1236

RESULT 11
US-09-614-221A-404
/ Sequence 404, Application US/09614221A
/ Patent No. 6723837
/ GENERAL INFORMATION:
/ APPLICANT: Karunananda, Balasubojini
/ APPLICANT: Yu, Jaehyuk
/ APPLICANT: Kishore, Ganesh M.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
/ TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
/ FILE REFERENCE: 16516.075
/ CURRENT APPLICATION NUMBER: US/09/614,221A
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/142,981
/ PRIOR FILING DATE: 1999-07-12
/ NUMBER OF SEQ ID NOS: 626
/ SEQ ID NO 404
/ LENGTH: 2964
/ TYPE: DNA
/ ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-404

Alignment Scores:
Pred. No.: 7.82e-19      Length: 2964
Score: 246.50           Matches: 103
Percent Similarity: 39.68%      Conservative: 68
Best Local Similarity: 23.90%    Mismatch: 135
Query Match: 10.79%            Indels: 125
DB: 4                    Gaps: 19

US-10-701-200-6 (1-437) x US-09-614-221A-404 (1-2964)

QY      31  LysProIlyshYsValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAla 50
Db      610  AAGAAGAAGAAGATGCTGTCATGACCTTCGTCGTCATTCCTCCAGGTAAAGATCCGCT 669
QY      51  IleGlySerLeuIleGlyuArgTYrThGluIleAspProSerIleGluIleIleCysTYr 70
Db      670  GTTCGTCGCCGTCGTCGTCGT-----ACAGGTATCATTTCCGCTGTCATGATGTTTGGCTGTT 723
QY      71  ArgGlyIlyTYrIlysgIlyLeuLeuGlyAspSerTYrProValThrAlaGluValArg 90
Db      724  TACGAAGGTAAACGAAGGTTACTTAAGAGCGCGTAAATAT-----TTAAG 768
QY      91  LysIlysAla-----GlyValLeuGlnArgPheGlyGlySerValIleGly 105
Db      769  AAAATGCTTCGGAAGATGTCAGAGCGTTGGTTAAGTAA---GGTGTACTTTCGATTCGT 825
QY      106  AsnSerArgValIlysLeuThrAsnValIlyAspCysValIlysArgGlyLeuValIysGlu 125
Db      826  ACTGCT-----CGTTCATGAATTCAGA 849
QY      126  GlyIlyAspProGlnIlysValAlaIleAspGlnLeuValIlyAspGlyValIleAspIleLeu 145
Db      850  AAGCGTAGGGTCCTGAACAAGCTGCAGGCAATTAAATTCGCAAGGATTAAGACGCTTGG 909
QY      146  HisThrIleGlyIlyAspAspThrAsnThrAlaIleAlaIleAspLeu----- 160
Db      910  GTTGTGTTGGTGGTGGATGCTTTTAACC---GGTCTGATCTTTTCAGACAGAAATGG 966
QY      161  -----AlaIlePheLeuAlaArgAsnAsnTYr----- 169
Db      967  CCATCTTTGGTGTATGAATGGCTTGCAGAGAGGTATGATCTAATAAGAAAGATGCGCCCA 1026
QY      170  -----GlyLeuThrValIleGlyLeuProIlysThrValIleAsnAspValPheProIle 187
Db      1027  TACAAGAATTTGTCATCTGTTGCTTCGTCGATCCATCATATGATGATGTCGTACT 1086
QY      188  LysGlnSerLeuGlyAlaTPThrAlaIleGluGlnIlyAlaArgTYrPheMetAsnVal 207
Db      1087  GACTCTACCATGGTGGCTTATCTGCTTTGGAAAGAACTGTGTAATAGTTGACTACATTT 1146

```


DB 1443 GCGGTGTCGGCTGCGCGCTGCGGGAAGCGCGCTGCGGACCTGCTGCGCCGCGACATC 1384
QY 313 -----GluValProArgAspAlaPheGlyHisIleLeu----- 323
DB 1389 GAGGGGAGATTCGGGTGACGGTGTGGGCGACCTCCACCGCGGGGCGAGCCCGCGCG 1324
QY 324 -----LeuAspAlaValAsnProGly 330
DB 1323 GCGGACCGGTACTGCGCCTGCTACGCGCTCAAGGTGTGAGCTGTGAGCGCGGCG 1264
QY 331 LysTrp 332
DB 1263 CAGTGG 1258

RESULT 13
US-08-956-171E-312
Sequence 312, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunach
GIL H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 312:
SEQUENCE CHARACTERISTICS:
LENGTH: 7972 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 312:
US-08-956-171E-312

Alignment Scores:
Pred. No.: 6,18e-16 Length: 7972
Score: 228.00 Matches: 102
Percent Similarity: 42.24% Conservative: 75
Best Local Similarity: 24.34% Mismatches: 147
Query Match: 9.98% Indels: 95
DB: 4 Gaps: 19

US-10-701-200-6 (1-437) x US-08-956-171E-312 (1-7972)
QY 26 TyrThrLeuMetAsnLysProLysLeuValAlaIleuThrAlaGlyLeuAlaPro 45
DB 3647 TATGCGTCATG-----AAGAAATTCAGATTAACTAGTGGGAGATTCACCT 3697
QY 46 CysLeuAsnSerAlaIleGlySerLeuIleGluArgLysThrGluIleAspProSerIle 65
DB 3698 GGAATGAATGCTGCGGTAAAGAGCAATTGTCGT-----ACGCAATTTTCAATGAATTT 3751
QY 66 GluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuGlyAspSerLysProVal 85
DB 3752 GAGGTTATGCTGCTATCATGCTTATCCAGAGTTGTTAAATGATGATATTCATAACTT 3811
QY 86 ThrAlaGluValArgLysLeuAlaGlyValLeuGlnArgPheGlyGlySerValIleGly 105
DB 3812 -----GAATTAGATCATGTTGGGATGATGATTCAGCCT---GGAGTTACATTCCTGTAT 3862
QY 106 AsnSerArgValLysLeuThrAsnValLysAspCysValLysArgGlyLeuValLysGlu 125
DB 3863 TCAGCAAGA-----TGTCCAGAG-----TTTAAGAG 3889
QY 126 GlyGluAspProGlnLysValAlaAlaAspGlnLeuValLysAspGlyValAspIleLeu 145
DB 3890 ---CAGAGATACGTAAAGTTGCAATCGAAACTTAAAGAGGATTCAGGCGCTT 3946
QY 146 HisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeuAlaAlaPheLeuAla 165
DB 3947 GTAGTTATGGTGGTGGAGCGATGATTCGCGGTGCAACAGCATGAGAGATGTAA 4006
QY 166 ArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrValAspAsnAspValPhe 185
DB 4007 GAA-----ATTCAAACTATCGTATTCCTGTGACATGACATGACATGATCAT 4054
QY 186 ProIleLysGlnSerLeuGlyAlaThrThrAlaAlaGlnGlnGlyAlaArgTyrPheMet 205
DB 4055 GGTACTGATTTTACATTCATTCGATTCGACACAGCATTAATACGATTAATGCTTACGAC 4114
QY 206 AsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHisGluValMetGly 225
DB 4115 AAATTTAGATGACTGCGTCATAGTCACGACGACGACATTTATCATT---GAAGCAATGGGC 4171
QY 226 ArgAsnGlyGlyTyrPheLeuThrAlaAlaThrAlaGlnGlnLysTyrArgLysLeuAspArg 245
DB 4172 CGTGAATGTGGA-----GATCTTA 4189
QY 246 AlaGluTyrPheLysProGlnLeuGlyLeuThrArgGlnSerTyrGluValHisAlaValPhe 265
DB 4190 GCATTATGG-----GCTGGATTA-----TCAGTTGGTGGCTGAGCAATGTGA 4231
QY 266 ValProGlnMetAlaIleAspLeuGlnAlaGluAlaLysArgLeuArgGluValMetAsp 285
DB 4232 GTTCAGAGAGTGAAGAAACAGAT-----ATTAAGAAATTAAGTCTGAT 4270
QY 286 LysValAspCysValAsnIlePheValSerGlnGlyAlaGlyValGlu----- 301
DB 4271 AAAATTGAA-----CAAGGATTTAAACGTGTGAAGAA 4303
QY 302 -----AlaIleValAlaGlnMetGlnAlaLysGlnGlnGluValProArgAspAla 318
DB 4304 CACTCAATCGTCTTGTGACGAAAGGTGTATGACTGCGCAAGATTCACAAAATAATTA 4363
QY 319 PheGlyHisIleLysLeuAspAlaValAsnProGlyLysTrpPheGly-----GluGln 336
DB 4364 TCACAATACATCAATGTTGAT-----AATAGAGTGTGCTTATGATTCACAGCTTCAACGT 4417
QY 337 PheAlaGlnMetIleGlyAlaGlnLysThrLeuValGlnLysSerGlyTyrPheAlaArg 356
DB 4418 GGTGTGACCCCAACAGCGTGGATGAGATTTTATGATCAGTCGTTTATGGTGAATATGG--- 4474
QY 357 AlaSerAlaSerAsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaVal 376
DB 4475 -----GTAGACTTATTAATG 4489

QY 377 GluCyAlaPheArgGluSerGlyValIleGlyHisAspGluSapAsnVal 396
 DB 4490 CAAGGTGAACAGCTAAGGCTGTGAAT-----AAGAACATTA 4531
 QY 397 LeuArgAlaIleGluPheProArgIleGlyGlyProPheAsnIleAspThrAsp 416
 DB 4532 ATTGTAGCAACATCTTTGATGAATTTTGTGATGAAGATCAATAATTGATTATAGT 4591
 QY 417 -TTPheAsnSerMetLeuSerGluIleGlyGlnProGlyGlyValGlu 434
 DB 4592 CTATATGAACTTCTGTAACAGATTATATTAAGATTTCAGAGGAATTATMAA 4646
 RESULT 14
 US-08-781-986A-312
 ; Sequence 312, Application US/08781986A
 ; Patent No. 6737248
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781, 986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; REFERENCE/DOCKET NUMBER: PB248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 312:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7972 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-781-986A-312
 Alignment Scores:
 Pred. No.: 6.18e-16 Length: 7972
 Score: 228.00 Matches: 102
 Percent Similarity: 42.24% Conservative: 75
 Best Local Similarity: 24.34% Mismatches: 147
 Query Match: 9.98% Indels: 95
 DB: 4 Gaps: 19
 US-10-701-200-6 (1-437) x US-08-781-986A-312 (1-7972)
 QY 26 TyrThrLeuMetAsnValProGlyValAlaIleLeuThrAlaGlyGlyLeuAlaPro 45
 DB 3647 TATGTCTGCTATG-----AAGAAATTCAGATTTAACCTAGTGAGATTCACT 3697
 QY 46 CylLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGluIleAspProSerIle 65
 DB 3698 GGAATGAAATCTGCTGTAAGACGATTGCTG-----ACAGCAATTTCAGTAAT 3751

QY 66 GluIleIleCysTyrArgGlyTyrGlyValLeuLeuGluValAspSerTyrProVal 85
 DB 3752 GAAGTTATGCTGTATCATGCTTACCAAGATTGTAAATGATGATATTAACCT 3811
 QY 86 ThrAlaGluValArgIleValAlaGlyValLeuGluArgPheGlyIleValIleGly 105
 DB 3812 -----GAATTAGATCATGAGTTGGGATACGATTCACGCT--GAGGTACATCTGTAT 3862
 QY 106 AsnSerArgValIleValThrAsnValIleAspCysValIleValGlyValGlu 125
 DB 3863 TCACGAGA-----TGTCCAGAG-----TTTAAGAG 3889
 QY 126 GluIleAspProGlnValAlaAlaAspGlnLeuValIleAspGlyValAlaPheLeu 145
 DB 3890 ---CAAGAGTACATTAAGATTGCAATCAAAAACCTACGTAAAGAGGATTGAGGCTT 3946
 QY 146 HisThrIleGlyIleAspPheThrAsnThrAlaAlaAlaAspLeuAlaPheLeuAla 165
 DB 3947 GTAGTTATGTTGTGACGATGATTCGCGGTGACACAGCATCATGAGAACTTAA 4006
 QY 166 ArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLeuThrValAspAsnAspValPhe 185
 DB 4007 GAA-----ATCAACTATCGGATTCCTCGTACGATTCGATCAATATCAAT 4054
 QY 186 ProIleValSerLeuGlyValAlaTyrThrAlaAlaGluGlnGlyValArgTyrPheMet 205
 DB 4055 GGTACTGATTTTCAATTTGATTTGACACAGCATTAATTCGATTATGCTTATGTCAC 4114
 QY 206 AsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHisGluValMetGly 225
 DB 4115 AAATTAGAGATCTGCTGATCACTGACGACGACAACTTATCACTT--GAAGCAATGCGC 4171
 QY 226 ArgAsnCysGlyTyrLeuThrAlaAlaThrAlaGlnGluTyrArgIleLeuAspArg 245
 DB 4172 CGGATTTGCA-----GATCTA 4189
 QY 246 AlaGluTyrLeuProGluLeuGlyLeuThrArgIleSerTyrGluValHisAlaValPhe 265
 DB 4190 GCATTATG-----GCTGATTA-----TCAGTTGCTGTCGACCAATGTA 4231
 QY 266 ValProGluMetAlaIleAspLeuGluAlaGluAlaValArgLeuArgIleValMetAsp 285
 DB 4232 GTTCCAGAGGTGAACAGAT-----ATTAAAGAAATGCTGAT 4270
 QY 286 IysValAspCysValAsnIlePheValSerGluGlyAlaGlyValGlu----- 301
 DB 4271 AAAATTGAA-----CAAGTATTAACGCTTAAGAA 4303
 QY 302 -----AlaIleValAlaGluMetGlnAlaIleGlyGlnGluValProArgAspAla 318
 DB 4304 CACTCAATGCTTCTGTAGCAGAGGTTGTATGATCGCCGCAATTTGCTAAAAAAGATT 4363
 QY 319 PheGlyHisIleIleValAspAlaValAlaAsnProGlyValTyrPheGly-----GluGln 336
 DB 4364 TCACAAATACATCAATGTGAT-----AATAGTGTCTGTGTAGTCACTCAACGT 4417
 QY 337 PheAlaGluMetIleGlyValGluValThrLeuValGlnIleValSerGlyTyrPheAlaArg 356
 DB 4418 GGTGTAGCCCAACAGCTGCGGATGATGTTTGATGCACTACGTTAGGTGATATGCG-- 4474
 QY 357 AlaSerAlaSerAsnValAspAspMetArgLeuIleValSerCysAlaAspLeuAlaVal 376
 DB 4475 -----GTAGCTTATTATG 4489
 QY 377 GluCyAlaPheArgGluSerGlyValIleGlyHisAspGluSapAsnVal 396
 DB 4490 CAAGGTGAACAGCTAAGGCTGTGAAT-----AAGAACATTA 4531
 QY 397 LeuArgAlaIleGluPheProArgIleGlyGlyProPheAsnIleAspThrAsp 416
 DB 4532 ATTGTAGCAACATCTTTGATGAATTTTGTGATGAAGATCAATAATTGATTATAGT 4591
 QY 417 -TTPheAsnSerMetLeuSerGluIleGlyGlnProGlyGlyValGlu 434

```

Db      4592 CTRATGAACTGCTACAGATTATCTATATTAAGATTTCAGGAGGAAATTATATAA 4646
RESULT 15
US-09-489-039A-5087
; Sequence 5087, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5087
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5087

Alignment Scores:
Pred. No.:      3,166-17      Length:      1017
Score:          227.00      Matches:      105
Percent Similarity: 38.84%      Conservative: 62
Best Local Similarity: 24.42%      Mismatches: 136
Query Match:      9.94%      Indels:      128
DB:               4          Gaps:      18

US-10-701-200-6 (1-437) x US-09-489-039A-5087 (1-1017)
QY      13 1LeaTgphCyehiSrTrpPhePheLeuAenPheAenPheTyTrhLeuMetAsnLySPro 32
Db      22 ATTATTTTTC-----TTCTAAGTCAGAGTAGCATGAT-----60
QY      33 LysLysValAlaIleLeuThrAlaGlyLysValProCysLeuAenSerAlaIleGly 52
Db      61 AAGAAATGCGGTGTGTGACAAAGTGGGTGATGCGCGGCGCATGAAACGACCAATTCGT 120
QY      53 SerLeuIleGluArgTyTrhGluLeuAenProSerIleGluIleIleCysTyTrhArgLy 72
Db      121 GCGGTGTGCGC-----GCGGCAATTAACGGAAGCTCGAAAGTTTGGAACTCTATGAC 174
QY      73 GlyTyTrpGlyLeu-----LeuLeuGlyAspSerTyTrpProValThrAla 87
Db      175 GGTACCTCGGATTTGATGAAAGACCGTATGCTCACTCGACCGCTTACAGCGTTTCCGAC 234
QY      88 GluValArgLysAlaGlyValLeuGlnArgPheGlyLysSerValIleGlyAenSer 107
Db      235 ATGATCAACCGC-----GGCGGTAACCTTCTCGGCTCGCCT 270
QY      108 ArgValLysLeuThrAenValLysAspCysValLysArgGlyLeuValLysGluGlyGlu 127
Db      271 CGC-----TTCCCGGAATTCGCGA-----GAA 294
QY      128 AspProGlnLysValAlaIleAspGlnLeuValLysAspGlyValAspIleLeuHisThr 147
Db      295 CACATCGCGGTGTGCTATCGAAACATGAAAGACGCGGCTGACCGCGCTGTG 354
QY      148 11eGlyGlyAspAspThrAenThrAlaAlaIleAspLeuAlaIlePheLeuAlaArgAen 167
Db      355 ATCGGCGGTGACGCTTCTATATGAGGCGATGCGCCTG-----ACC 396
QY      168 AenTyTrpGlyLeuThrValIleGlyLeuProLysTrpValAspAsnAspValPheProIle 187
Db      397 GAGATGGGCTTCACATGCAATCGGCTGCGGCGCACCATGACATATCAAGGCACT 456
QY      188 LysGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGlyAlaArgTyTrpMetAenVal 207
Db      457 GACTACACATCGGCTTCTTACTGCTGAGACCGGTGTGGAACGATTCACGCTTTG 516

```

```

QY      208 ValAlaGluAenAsnAlaAsnProArgMetLeuIleValIleGluValMetGlyArgAen 227
Db      517 CGCGACACCTCTCTTCGCACACGATATCTCCGTGTG---GAAGTGTGGGCGGTTAC 573
QY      228 CysGlyTrpLeuThrAlaAlaThrAlaGlnGlyTrpArgLysLeuLeuAspArgAlaGlu 247
Db      574 TGTGGCACTGACCTGCGCGCGCGC-----ATTGCGCGCGCGGTGTAG 618
QY      248 Tripleu-----ProGluLeuGlyLeuThrArgGlnSerTyTrpGluValHisAlaValPhe 265
Db      619 TTCATCATGCTGCTCGAAGTGAATATACCGCTGAC-----654
QY      266 ValProGluMetAlaIleAspLeuGluAlaGlyValLys-----278
Db      655 -----GATCTGGTGCAGAAATCAAAAGCCGGTATCGCGAAAGGAAA 696
QY      279 -----ArgLeuArgGluValMetAspLysValAspCysValAenIlePhe 293
Db      697 AAACACGCTATCGTGGCCATCACGACGACATGTGCGACGTTGACGAGCTGGCAAGCTAC 756
QY      294 ValSerGluGlyAlaGlyValGlu-----AlaIleValAlaGluMetGlnAlaLys 310
Db      757 ATCGAGAAAGAGACTGCGCGTGAACCTGCGGCACCGTCTCGGCGCACATTCACGCGCGC 816
QY      311 GlyGlnGluValProArgAspAlaPheGlyHisIleLysLeuAspAlaValAsnProGly 330
Db      817 GGTTCCTCCGTTCTTACGAC-----837
QY      331 LysTrpPheGlyGluGlnPheAlaGlnMetIleGlyAlaGluLysTrpLeuValGlnLys 350
Db      837 -----837
QY      351 SerGlyTyTrpPheAlaArgAlaSerAlaSerAenValAspAspMetLysLeuIleLysSe 370
Db      838 -----CGCATCTGGCTTCCCGCATGCGCGGCGCTTATGCGATGAGCTGCTG 882
QY      370 rCysAlaAspLeuAlaValGluCysAlaPheArgArgGlnSerGlyValIleGlyHisAs 390
Db      883 CTGCAGGCGCCATGGCGGC- CGTTGCGTGGCATTCAGAAACGAAAGCTGTGACACACGA 941
QY      390 pGluAspAenGlnAenValLeuArgAlaIleGluPheProArgIleLysGlyLysPr 410
Db      942 C-----ATCATGATGCCATCGAAGAACATGAAG-----CGTCC 974
QY      410 PheAsnIleAspThrAspTrpPheAsn 419
Db      975 GTTCAAG-----AACGACTGCTGAT 996

```

Search completed: March 1, 2005, 17:14:53
 Job time : 265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 17:01:32 ; Search time 749 Seconds
(without alignments)
3457.177 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVVTMPHYLTADIRCFWFF.....FNSMSEIGQPKGKVEVSH 437

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789605

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-DB=PubMed Applications NA -QFMT=faaCap SURF1=rmph -MIMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=500 -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEADSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USRR=US10701200 @CGN 1.1480 @runat 25022005 102834_10410
-NCPU=6 -ICPU=3 -NO_MMP -LARGEQUERY -NEG_SCORE=0 -WAIT -DSBIOCK=100
-LONGLOG -DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FAPOP=6 -FAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : 1: Published Applications NA:*

2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|---------|----|----------------------|---------------------|
| 1 | 2284 | 100.0 | 1311 | 9 | US-09-934-901-15 | Sequence 15, Appl 1 |
| 2 | 2284 | 100.0 | 1311 | 9 | US-09-934-868-5 | Sequence 5, Appl 1 |
| 3 | 2284 | 100.0 | 1311 | 10 | US-09-941-947A-1 | Sequence 1, Appl 1 |
| 4 | 2284 | 100.0 | 1311 | 15 | US-10-320-924-15 | Sequence 15, Appl 1 |
| 5 | 2284 | 100.0 | 1311 | 15 | US-10-320-874-15 | Sequence 1, Appl 1 |
| 6 | 2284 | 100.0 | 1311 | 17 | US-10-363-567-1 | Sequence 15, Appl 1 |
| 7 | 2284 | 100.0 | 1311 | 18 | US-10-321-210-15 | Sequence 15, Appl 1 |
| 8 | 659 | 28.9 | 552 | 17 | US-10-368-493-38035 | Sequence 38035, A |
| 9 | 306.5 | 13.4 | 1050 | 17 | US-10-368-493-31795 | Sequence 31795, A |
| 10 | 298 | 13.0 | 9025608 | 15 | US-10-156-761-1 | Sequence 1, Appl 1 |
| 11 | 294 | 12.9 | 1026 | 15 | US-10-156-761-6060 | Sequence 6060, Ap |
| 12 | 293 | 12.8 | 1570 | 17 | US-10-425-114-9204 | Sequence 9204, Ap |
| 13 | 293 | 12.8 | 1765 | 17 | US-10-425-114-6293 | Sequence 6293, Ap |
| 14 | 293 | 12.8 | 1945 | 17 | US-10-425-114-10982 | Sequence 10982, A |
| 15 | 293 | 12.8 | 4413 | 17 | US-10-424-599-141147 | Sequence 141147, A |
| 16 | 284.5 | 12.5 | 1023 | 15 | US-10-156-761-7102 | Sequence 7102, Ap |
| 17 | 284.5 | 12.5 | 9025608 | 15 | US-10-156-761-1 | Sequence 1, Appl 1 |
| 18 | 279.5 | 12.2 | 1998 | 18 | US-10-437-963-101981 | Sequence 101981, Ap |
| 19 | 270 | 11.8 | 1347 | 17 | US-10-282-122A-11067 | Sequence 11067, A |
| 20 | 269.5 | 11.8 | 1023 | 15 | US-10-156-761-2809 | Sequence 2809, Ap |
| 21 | 269.5 | 11.8 | 2100 | 18 | US-10-767-701-14854 | Sequence 14854, A |
| 22 | 267.5 | 11.7 | 3405 | 18 | US-10-425-115-165105 | Sequence 165105, A |
| 23 | 263.5 | 11.5 | 2310 | 18 | US-10-437-963-12274 | Sequence 12274, A |
| 24 | 262 | 11.5 | 960 | 17 | US-10-369-493-26592 | Sequence 26592, A |
| 25 | 261 | 11.4 | 1035 | 17 | US-10-369-493-42948 | Sequence 42948, A |
| 26 | 259 | 11.3 | 1358 | 18 | US-10-437-963-6352 | Sequence 6352, Ap |
| 27 | 258.5 | 11.3 | 2663 | 18 | US-10-425-115-121463 | Sequence 121463, A |
| 28 | 255 | 11.2 | 1677 | 17 | US-10-425-114-8526 | Sequence 8526, Ap |
| 29 | 254 | 11.1 | 2149 | 17 | US-10-424-599-128266 | Sequence 128266, A |
| 30 | 249.5 | 10.9 | 956 | 17 | US-10-369-493-33339 | Sequence 33339, A |
| 31 | 248 | 10.9 | 2123 | 18 | US-10-425-115-119704 | Sequence 119704, A |
| 32 | 246.5 | 10.8 | 1073 | 17 | US-10-368-493-33356 | Sequence 33356, A |
| 33 | 246.5 | 10.8 | 2664 | 17 | US-10-369-493-45685 | Sequence 45685, A |
| 34 | 246.5 | 10.8 | 2364 | 18 | US-10-793-639-404 | Sequence 404, Ap |
| 35 | 242.5 | 10.6 | 2024 | 17 | US-10-424-599-68225 | Sequence 68225, A |
| 36 | 242.5 | 10.6 | 2024 | 17 | US-10-425-114-11330 | Sequence 11330, A |
| 37 | 241.5 | 10.6 | 1441 | 17 | US-10-425-114-10007 | Sequence 10007, A |
| 38 | 241.5 | 10.6 | 2812 | 17 | US-10-424-599-115967 | Sequence 115967, A |
| 39 | 240.5 | 10.5 | 960 | 17 | US-10-282-122A-36804 | Sequence 36804, A |
| 40 | 240.5 | 10.5 | 963 | 9 | US-09-815-242-9687 | Sequence 9687, Ap |
| 41 | 240.5 | 10.5 | 963 | 17 | US-10-282-122A-38849 | Sequence 38849, A |
| 42 | 240.5 | 10.5 | 963 | 17 | US-10-282-122A-39841 | Sequence 39841, A |
| 43 | 239.5 | 10.5 | 3661 | 18 | US-10-437-963-1590 | Sequence 1590, Ap |
| 44 | 238.5 | 10.4 | 963 | 9 | US-09-815-242-6325 | Sequence 6325, Ap |
| 45 | 238.5 | 10.4 | 963 | 17 | US-10-369-493-47325 | Sequence 47325, A |

ALIGNMENTS

RESULT 1
US-09-934-901-15
; Sequence 15, Application US/09934901
; Patent No. US20020110885A1
; GENERAL INFORMATION:
; APPLICANT: Kofifae, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: No. US20020110885A1ton, Kelley C.
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C1619 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 906
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office. 97
; SEQ ID NO 15
; LENGTH: 1311
; TYPE: DNA

ORGANISM: METHYLOMONAS SP.
US-09-934-901-15

Alignment Scores:

| Pred. No.: | 4,19e-253 | Length: | 1311 |
|------------------------|-----------|---------------|------|
| Score: | 2284.00 | Matches: | 437 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-10-701-200-6 (1-437) x US-09-934-901-15 (1-1311)

```

QY 1 AapValValThrrProcyrHsleuthrAlaapRlleargPheCyehistrPhephe 20
DB 1 GATGGGTGACATGGCGCTTATCACTTAAGCGCTGATTAATTCGATTGTCATGGTTTTT 60
QY 21 LeuAapPheAapPheYrThrLeuMeAapPheAapPheAapPheAapPheAapPheAapPhe 40
DB 61 CTTAACTTTAACTTTCAACGCTCATGAACAACTTAATAAAAGTTGCAATCTGACAGCA 120
QY 41 G1yG1yLeuAlaProCyAeAapPheAapPheAapPheAapPheAapPheAapPheAapPhe 60
DB 121 GCGGCGTGGCGCTTGTGAATTCGCAATCGGTAATGTTGATCGAATGATTAACCGAA 180
QY 61 IleAapProserIleG1uIleIleCyeryrArgG1yYrIlyG1yLeuLeuLeuG1y 80
DB 181 ATCGATCTTACGATGAATGAATCATTTGCTATCGCGGCGTTAATAAGCGCTGTGCGGC 240
QY 81 AapSerYrProValThrrAlaGluValArgLysValaG1yValLeuG1naRppheG1y 100
DB 241 GATTCTTATCCAGTAAAGCGCCGAAAGCGCTTAATAAAAGCGGCTTCTGCAACGTTTGGC 300
QY 101 G1ySerValIleG1yAapSerAargValIyLeuThraAapValIyAapCyValIyAarg 120
DB 301 GGTTCGTGATCGGCAACGCGCGCTCAATTAAGCAATGTCATAAGCTGCGTAAGC 360
QY 121 G1yLeuValIyG1uG1yG1uAapProG1uIyValIleAlaAapG1uLeuValIyAap 140
DB 361 GGTGGTGTAAAGAGGGTGAATCGCAAAAGTCGCGGCTGATCAATTTGTTAAGGAT 420
QY 141 G1yValAapIleAapIleAapIleAapIleAapIleAapIleAapIleAapIleAapIle 160
DB 421 GGTGTCGATTTCTGCAACCATCGCGCGCTGATTAACCAATACCGACAGCGATTTG 480
QY 161 AlaAlaPheLeuAlaAargAapAapIyG1yLeuThraValIleG1yLeuProIyAap 180
DB 481 GCAGCATCTCTGCGCAAGATATTAATTAAGCTGACGCTCATTTGTTAATAAACCGTC 540
QY 181 AapAapAapValPheProIleIyG1uAapIleAapIleAapIleAapIleAapIleAapIle 200
DB 541 GATTAAGAGATTTTCCGATCAAGCATCACTAAGGTGCTTGAATGCGCCGCAAGCAAGC 600
QY 201 AlaArgYrPheMeAapValIyAlaGluAapAapAapAapAapAapAapAapAapAapAap 220
DB 601 GCGGCTTATTTATGATGACGTGCGCGCAAAACCAAGCCCAACGCTGATTCGTA 660
QY 221 H1eG1uValIyMeG1yAargAapCyG1yYrPheuthrAlaAlaThraIaG1uIyAarg 240
DB 661 CACGAAGTATGGCGCGTAACTGCGGCTGCGAAGCGCTGCAACCGCGCAAGATATGCG 720
QY 241 LysLeuLeuAapAargAlaG1uIyTrPleuPProG1uLeuG1yLeuThraArgIySerYrG1u 260
DB 721 AAATTAATCTGGACCGTCCAGTGGTGGCGGAATGGGTTTGACTCGTAATCTTAATGAA 780
QY 261 ValH1eAlaValPheValProG1uMeAlaIleAapLeuG1uAlaG1uAlaIyAargLeu 280
DB 781 GTGCACGCGGATTTGCTTCCGAATGCGGATCGACTGAAGCGCAAGCGCAAGCGCGCTG 840
QY 281 ArgG1uValIyMeAapIyAapValAapCyValAapIleAapValSerG1uIyAlaG1yVal 300
DB 841 CCGGAAGTATGACAAAGTCAATTCGCTCAACATCTTCTGTTCCGAAGGTCGCGGCGTC 900

```

```

QY 301 G1uAlaIleValAlaG1uMeG1naIyAapIleAapIleAapIleAapIleAapIleAapIle 320
DB 901 GAGCTATCGTCGCGGAATGAGGCGCAAGCGCAAGGTCGCGGATGCTTGGC 960
QY 321 H1eIleLeuAapAapValAapProG1yYrTrPheG1yG1uG1uPheAlaIleMe 340
DB 961 CACATCAATCGATGCGGCTCAACCTGTAATGTTGCGGAGCAATTCGCGCAAGT 1020
QY 341 I1eG1yAlaG1uIyYrThrLeuValG1uIyAapSerG1yYrPheAlaAargAlaSer 360
DB 1021 ATAGCGCGGAAAAAACCTGCTACAAAAATCGGATCTTTCGCGGCTTCTCTCC 1080
QY 361 AapValAapAapAapAapAapAapAapAapAapAapAapAapAapAapAapAapAapAap 380
DB 1081 AACGTTGACGACATCGCTTATCAATCGTGGCGCACTTGGCGTCAAGTCCGCTTC 1140
QY 381 ArgArgIySerG1yValIleG1yH1eAapG1uAapAapAapAapAapAapAapAapAap 400
DB 1141 CCGCGCGAGTCTGGGCTATCGGTCACGACGAAGCAACGCGCAAGCTGTTGCTCGATC 1200
QY 401 GluPheProArgIleIyG1yYrPheAapIleAapThraAapTrPheAapSer 420
DB 1201 GAGTTCCGCGCATCAAGGCGCGCAACGTTCAATATCGACACCGACTGTTCAATAGC 1260
QY 421 MetLeuSerG1uIleG1yG1uProIyG1yG1yValG1uValSerHis 437
DB 1261 ATGTTGACGAATCGCGCAAGCTTAAGCGGTAAAGTCGAAGTCAGCCAC 1311

RESULT 2
US-09-934-868-5
; Sequence 5, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Kofas, Matheos
; APPLICANT: Odem, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C11596 US NA
; CURRENT APPLICATION NUMBER: US/09/934, 868
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229, 858
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: METHYLOMONAS SP.
US-09-934-868-5

Alignment Scores:
Pred. No.: 4,19e-253 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-701-200-6 (1-437) x US-09-934-868-5 (1-1311)
QY 1 AapValValThrrProcyrHsleuthrAlaapRlleargPheCyehistrPhephe 20
DB 1 GATGGGTGACATGGCGCTTATCACTTAAGCGCTGATTAATTCGATTGTCATGGTTTTT 60
QY 21 LeuAapPheAapPheYrThrLeuMeAapPheAapPheAapPheAapPheAapPheAapPhe 40
DB 61 CTTAACTTTAACTTTCAACGCTCATGAACAACTTAATAAAAGTTGCAATCTGACAGCA 120
QY 41 G1yG1yLeuAlaProCyAeAapPheAapPheAapPheAapPheAapPheAapPheAapPhe 60
DB 121 GCGGCGTGGCGCTTGTGAATTCGCAATCGGTAATGTTGATCGAATGATTAACCGAA 180
QY 61 IleAapProserIleG1uIleIleCyeryrArgG1yYrIlyG1yLeuLeuLeuG1y 80

```

```

DB      181 ATCGATCTTACGTAATAATCATTTGCTATCGCGCGGTTATTAAGCCCTGTTGCGGCG
QY      81 AAPPSErTYrProValThrAlaGluValArgLysLysAlaGlyValLeuGlnArgPheGly 100
DB      241 GATTCTTATTCGATTAACGGCCGAGAGTGGCTTAAGAGCGGGGTCTTCTGCAACCTTTTGGC 300
QY      101 GlysSerValIleGlyAnsSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
DB      301 GGTTCTGTGATCGGCAACAGCCGCGTCAATTACCAATGTCAAGACTGCGTGAACCG 360
QY      121 GlyLeuValLysGluGlyLysAspProGlnLysValAlaAlaAspGlnLeuValLysAsp 140
DB      361 GGTTTGTGCAAGAGGAGGTGAAGATCCGCAAAAAGTCGGCGCTGATCAATTGGTTAAGCAT 420
QY      141 GlyValAspIleLeuHisThrIleGlyLysAspAspThrAsnThrAlaAlaAspLeu 160
DB      421 GGTTGTGATATTTGACACCATCGCGCGGATGATACCAATACGGACGACGCGGATTTG 480
QY      161 AlaAlaPheLeuAlaArgAsnAsnTYrGlyLeuThrValIleGlyLeuProLysThrVal 180
DB      481 GCGACATTCCTGCGCGAATAATATTACGACCTGACCGTCACTTGTCTTACCTAACAACGTC 540
QY      181 AAPAsnAspValPheProIleLysGlnSerLeuGlyAlaIleThrAlaAlaGlnGlnGly 200
DB      541 GATTAACGAGTATTCGATCAACGATCACTAGTGTGCTTGACCTCGCGCGAAGCGC 600
QY      201 AlaArgTYrPheMetAsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
DB      601 GCGCGTATTTCTGTAACCGTGGCGCGCAAAACAGCCCAACCGACGATGCTGATCGTA 660
QY      221 HisGluValMetGlyArgAsnCysGlyTYrLeuThrAlaAlaThrAlaGlnGluTYrArg 240
DB      661 CACGAAGTGTGGCGCGTAACTCGCGGTGCTACCGCTGCAACCGCGCAAGAAATCCG 720
QY      241 LysLeuLeuAspArgAlaGluTYrLeuProGlnLeuGlyLeuThrArgGlnSerTYrGlu 260
DB      721 AATTATCTGACCGTCCGCGAATGCGGATTTGGCTTGACTCGTGAATCTTATGAA 780
QY      261 ValHisAlaValPheValProGlnMetAlaIleAspLeuGlnAlaGlnAlaLysArgLeu 280
DB      781 GTTCACGCGGATATTCCTCGGAAATGCGCATGACCTGGAACCGAAGCCAGCCCTG 840
QY      281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGlnGlyAlaGlyVal 300
DB      841 CCGGAAGTGTGACAAAGTCCGATTCGCTCAACATCTTCCCTTCCCAAGGTGCGCGCTC 960
QY      301 GlnAlaIleValAlaGlnMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
DB      901 GAAGCTATCGTCCGGAATGCAAGCCCAAGCCCAAGAGTGC CGCGCATGCGTCCGCG 960
QY      321 HisIleLysLeuAspAlaValAsnProGlyLysTYrPheGlyGlnPheAlaGlnMet 340
DB      961 CACATCAAACTGATCGGTCAACCTGTGTAATGTTCGCGCAAGCAATTCGCGCGCATG 1020
QY      341 IleGlyAlaGluTYrThrLeuValGlnLysSerGlyTYrPheAlaArgAlaSerAlaSer 360
DB      1021 ATTAGCGCGGAAACCCCTGGTACAAATCCGGATACCTTCCCGCTGCTTCTGCTTCC 1080
QY      361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaAlaGlnCysAlaPhe 380
DB      1081 AACGTGACGACATGCGTTGATCAATCGTGC CGGACTTGCGCGGTGCGCGCTTC 1140
QY      381 ArgArgGlnSerGlyValIleGlyLysAspGlnAspGlnLysValLeuAlaGlnAla 400
DB      1141 CGCGCGAGCTCGCGTATCGGTCAACGCAAGACAAACGCGCAACGTTGCTGCGCATC 1200
QY      401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspThrAspTYrPheAsnSer 420
DB      1201 GAGTTTCGCGGATCAAGGCGCGCAACCTTCAATATGCAACCGACTGTTCAATAGC 1260
QY      421 MetLeuSerGlnIleGlyGlnProLysGlyLysValGlnValSerHis 437

```

```

DB      1261 ATGTTGAGCGAAATCGGCCACGCTTAAGCGGTTAAAGTCGAAGTCAGCCAC 1311
RESULT 3
US-09-941-947A-1
; Sequence 1, Application US/09941947A
; Publication No. US20030003528A1
; GENERAL INFORMATION:
; APPLICANT: Brzoskiewicz, Patricia C.
; APPLICANT: Cheng, Qiong
; APPLICANT: DiCosimo, Deana J.
; APPLICANT: Kofias, Matheos
; APPLICANT: Miller, Edward S. Jr.
; APPLICANT: Odom, J. Martin
; APPLICANT: Picataggio, Steve
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
; FILE REFERENCE: CL1903 US NA
; CURRENT APPLICATION NUMBER: US/09/941,947A
; PRIOR FILING DATE: 2001-09-01
; PRIOR APPLICATION NUMBER: 60/229,907
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Methylobomonas 16a
US-09-941-947A-1

Alignment Scores:
Pred. No.: 4,19e-253 Length: 1311
Score: 2284.00 Matches: 437
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-701-200-6 (1-437) x US-09-941-947A-1 (1-1311)
QY      1 AAPValAlaThrTYrProTYrHisLeuThrAlaAspIleArgPheCysHisTYrPhePhe 20
DB      1 GATGTGTGACATGAGCCCTATCACTTAACGGCTGATATTCGATTTGTGATTTT 60
QY      21 LeuAsnPheAsnPheTYrThrLeuMetLeuLysProLysLysValAlaIleLeuThrAla 40
DB      61 CTTAACCTTAACCTTCAACAGCTCATGAACAAACCTTAAGAGTTGCAATTAAGACAGCA 120
QY      41 GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTYrThrGlu 60
DB      121 GCGCGCTTGGCGCTTGTGTAATCCGCATCGGATCGGTGATTCGAACGTTATACGAA 180
QY      61 IleAspProSerIleGluIleIleCysTYrArgGlyTYrLysGlyLeuLeuLeuGly 80
DB      181 ATCGATCTTACGTAATAATCATTTGCTATCGCGGCGTATTAAGCCCTGTTGCTGGCG 240
QY      81 AAPPSErTYrProValThrAlaGluValArgLysLysAlaGlyValLeuGlnArgPheGly 100
DB      241 GATTCTTATTCGATTAACGGCCGAGAGTGGCTTAAGAGCGGGGTCTTCTGCAACCTTTTGGC 300
QY      101 GlysSerValIleGlyAnsSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
DB      301 GGTTCTGTGATCGGCAACAGCCGCGTCAATTACCAATGTCAAGACTGCGTGAACCG 360
QY      121 GlyLeuValLysGluGlyLysAspProGlnLysValAlaAlaAspGlnLeuValLysAsp 140
DB      361 GGTTTGTGCAAGAGGAGGTGAAGATCCGCAAAAAGTCGGCGCTGATCAATTGGTTAAGCAT 420
QY      141 GlyValAspIleLeuHisThrIleGlyLysAspAspThrAsnThrAlaAlaAspLeu 160
DB      421 GGTTGTGATATTTGACACCATCGCGCGGATGATACCAATACGGACGACGCGGATTTG 480

```


| | | | |
|----|-----|---|-----|
| Db | 61 | CTTAACTTTAACTTCTACACGCTCATGACAAACCTTAAAAAGTTGGCAATCTACGACGA | 120 |
| Qy | 41 | GLYGLYLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGlu | 60 |
| Db | 121 | GGCGGCTTGCGGCTTGTTGAATTCGCGACATCGTAAGTTGATCGAACGTTATACCGAA | 180 |
| Qy | 61 | IleAspProSerIleGluIleIleCysTyrArgGlyTyrIleGlyLeuLeuGly | 80 |
| Db | 181 | ATGATCTCTAGCATAGAAATCATTTGGTATCGCGCGGTTATTAAGGCTGTTCTGGGC | 240 |
| Qy | 81 | AspSerTyrProValThrAlaGluValArgIleLeuValAspGlyValLeuGluArgPheGly | 100 |
| Db | 241 | GATTTCTTATCAGTACGACGCGCGAAGTGCATTAAGGCGGGGTGTTCTGCACAGTTTGGC | 300 |
| Qy | 101 | GlySerValIleGlyAsnSerArgValLeuThrAsnValIlyAspCysValIlyArg | 120 |
| Db | 301 | GGTTCTGTGATCGGCAACAGCCGGGTCAAAATTGCACAAATGTCAAAAGACTGGGTGAAACGC | 360 |
| Qy | 121 | GlyLeuValIlyGluGluGlyIlyAspProGluIlyValAlaIleAspGluLeuValIlyAsp | 140 |
| Db | 361 | GGTTTGGTCAAGAGGAGGTGAAGATCCCAAAAAGTCCCGGCTGATCAATTGGTTAAGAT | 420 |
| Qy | 141 | GlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaIleAspLeu | 160 |
| Db | 421 | GGTGTCATTTTCGCACACCATCGCGCGGATGATCCAAATACGCGACGACGCGGATTTG | 480 |
| Qy | 161 | AlaIlePheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProIlyThrVal | 180 |
| Db | 481 | GCAAGATTCCTGGCCGAAGAAATTAATTACGAGTACGCTCATTTGGTTTACCTAAACCGTC | 540 |
| Qy | 181 | AspAsnAspValPheProIleIlyGlnSerLeuGlyValATPThrAlaAlaGluGlnGly | 200 |

2.1. $\text{Pb}(\text{CH}_3\text{COO})_2 \cdot 3\text{H}_2\text{O}$ and $\text{Pb}(\text{CH}_3\text{COO})_2$

| | | | |
|----|------|---|------|
| Db | 61 | CTTAACTTTAACTTCTAACAGCGCTCATGAACAACTTAAATAAGTTGCAATATCGACAGCA | 120 |
| Qy | 41 | GlyGlyLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGlu | 60 |
| Db | 121 | GCGGCGCTTGGCGCGCTTGTTGAAATTCGCGCAATCGATGTTTGAATCGAAGCTTAAACCGAA | 180 |
| Qy | 61 | IleAspProSerIleGlyIleIleCysTyrArgGlyGlyTyrLeuGlyLeuLeuGly | 80 |
| Db | 181 | ATGATCTTACATGACATGAAATCATTTGCTATACGCGCGGTATTAAGAGCTGTCTGGCG | 240 |
| Qy | 81 | AspSerTyrProValThrAlaGluValArgLeysValIleGlyValLeuGlnArgPheGly | 100 |
| Db | 241 | GATTCTTATCATCAGTAAACGCCGGAAGTCGTTAAAGGCGGGTGTCTCGCAACGTTTGGC | 300 |
| Qy | 101 | GlySerValIleGlyAsnSerArgValLeysLeuThrAsnValIysAspCysValIlyArg | 120 |
| Db | 301 | GGTTCTGTGATCGGCACACCGCGGTCAAAATTGACCAATGTCAAAAGACTGCGTGAACGC | 360 |
| Qy | 121 | GlyLeuValIlySerGlyGlyIleAspProGlnIlyValAlaIleAspGlnIleValIlyAsp | 140 |
| Db | 361 | GGTTTGGTCAAGAGGGGTGAAGATCCGCAAAAGATCCGGCTGATCAATTGGTTAAGAT | 420 |
| Qy | 141 | GlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeu | 160 |
| Db | 421 | GGTGTGATATTTCTGCACACCATCGCGCGGATATATCAATACGCAACGCGGATTTG | 480 |
| Qy | 161 | AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProIysThrVal | 180 |
| Db | 481 | GCAGCATTTCTGGCGAGAAATTAATTAACGACTGACCGTCAATTGGTTTAACTTAAACCGTC | 540 |
| Qy | 181 | AspAsnAspValPheProIleIysGlnSerLeuGlyAlaThrThrAlaIleGluGlnGly | 200 |
| Db | 541 | GATTAACACATATTTCCGATCAAGCAATCATAGTGTCTTGACATGCGCGCGACCAAGGC | 600 |
| Qy | 201 | AlaArgTyrPheMetAsnValValAlaGluAsnAsnAlaAsnProAspMetLeuIleVal | 220 |
| Db | 601 | GCGGCTTATTTCATGAACGTGGTGGCCGAAACACGCCACCAACGATCTGATCGTA | 660 |
| Qy | 221 | HisGluValMetGlyArgAsnCysGlyTyrPheThrAlaAlaThrAlaGlnGlyTyrArg | 240 |
| Db | 661 | CACGAACTGATGGGCGGTAACTGCGCTGGGTGACCGCTGCAACGCGCGAGGAATATCGC | 720 |
| Qy | 241 | LysLeuLeuAspArgAlaGluTyrPheProGluLeuGlyLeuThrArgGlySerTyrGlu | 260 |
| Db | 721 | AAATTACTGACCGTCCGAGTGGTTCCCGAAATGGGTTGACTCGTGAATCTTAATGAA | 780 |
| Qy | 261 | ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaIlyArgLeu | 280 |
| Db | 781 | GTGACCGCGGTAATTCGTTCCGAAATGGCATTCACCTGGAAAGCCGAAAGCCAAACGCGCTG | 840 |
| Qy | 281 | ArgGluValMetAspIlyValAspCysValAsnIlePheValSerGluGlyAlaGlyVal | 300 |
| Db | 841 | CGGCAATGATGACAAAGTCGATGGTCAACATCTTCGTTTCCGAAGGTCCGGGGGCTC | 900 |
| Qy | 301 | GluAlaIleValAlaGluMetGlnAlaIlyGlyGlnGluValProArgAspAlaPheGly | 320 |
| Db | 901 | GAAAGCTATTCGTCGGGAAATGACAGGCCAAAGGCCAGAGAAAGTCCGCCGCTTCCGC | 960 |
| Qy | 321 | HisIleIysLeuAspAlaValAsnProGlyIlySerThrPheGlyGlyGlnPheAlaGluMet | 340 |
| Db | 961 | CACATCAAACTGATGGGTCAACACCTGTGTAAATGTTGGCGGAGCAATTCGCCACAGTG | 1020 |
| Qy | 341 | IleGlyAlaGluIysThrLeuValAlaGlnIlySerGlyTyrPheAlaArgAlaSerAlaSer | 360 |
| Db | 1021 | ATAGGCGCCGAAAAAACCTCTGTGTCAAAAATTCGGGATATCTTCGCCGCTTCTGCTTCC | 1080 |
| Qy | 361 | AsnValAspAspMetArgLeuIleIysSerCysAlaAspLeuAlaValGluCysAlaPhe | 380 |
| Db | 1081 | AAAGTTGACCAATGCGTTGATCAATTCGAGCGCCGACTTGGCGGTCTGAGTGGCGCTTC | 1140 |
| Qy | 381 | ArgArgGlySerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle | 400 |

Db 1141 CGCCGCACTCTGGGTATCGGTACAGCAAGACACCGCACTGTGGTGGATC 1200
Qy 401 GIupheProarGilelyeGlylybProPheasniLeasPThAsPThPheasner 420
Db 1201 GAGTTTCGGCGATCAAGCGGCAACCGTTCAATATGACACCGATCGTTCAATAGC 1260
Qy 421 MetLeuSerGluileGlyGlnProlyseGlylybValGluValSerHis 437
Db 1261 ATGTTGAGGAATCGCGCCAGCTAAAGCGGTAAAGTCGAAGTCAGCCAC 1311
RESULT 6
US-10-363-567-1
/ Sequence 1, Application US/10363567
/ Publication No. US20040077068A1
/ GENERAL INFORMATION:
/ APPLICANT: E. I. du Pont de Nemours & Company
/ TITLE OF INVENTION: CARBONENOID PRODUCTION FROM A SINGLE CARBON SOURCE
/ FILE REFERENCE: CL1903 PCT
/ CURRENT APPLICATION NUMBER: US/10/363,567
/ PRIOR FILING DATE: 2003-09-04
/ PRIOR APPLICATION NUMBER: US 60/229858
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: US 60/229907
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 1
/ LENGTH: 1311
/ TYPE: DNA
/ ORGANISM: Methylobionas 16a
US-10-363-567-1

Alignment Scores:

| Pred. No.: | 4,19e-253 | Length: | 1311 |
|------------------------|-----------|---------------|------|
| Score: | 2284.00 | Matches: | 437 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| | 17 | Gaps: | 0 |

US-10-701-200-6 (1-437) x US-10-363-567-1 (1-1311)

Qy 1 AsPValaThThTrPProGlyHileuThraAlaSPilaRPhCyHistrPhePhe 20
Db 1 GATGGGTGACATGCGCTCATCACTTAACGGCTGATATCGATTTGTGATCGTTT 60
Qy 21 LeuAnPheAsnPhetThrLeuMetAsnLysProlyseValAlaIleuThra 40
Db 61 CTTAACCTTAACCTTCTACAGCTCATGAACAACTAAAGAGTTGCAATCTGACAGCA 120
Qy 41 GilyGlyLeuAlaProCyLeuAsnSerAlaileGlySerLeuileGluArgTyThrglu 60
Db 121 GCGCGGCTTGGCGCTTGTGAATTCGCAATCGAGTTGATGCAAGTTATCCGA 180
Qy 61 IleAsPProSerileGluileileCyTyraGlyGlyTyTylyseGlyLeuLeuengly 80
Db 181 ATCGATCTGATGATGAATAATCTTGTCTATCGCGGGGTATTAAGCCCTGTGTGGCGC 240
Qy 81 AsPserTyProValThraIaGluValArgLysValaileGlyValLeuGlnArgPheGly 100
Db 241 GATTTCTTATCCAGTAACGGCGCAAGTGGTAAGAAAGCGGGGTTTGGCAAGTTTGGC 300
Qy 101 GilySerValileGlyAsnSerArgValLysLeuThraSnValLysAPCyValLysArg 120
Db 301 GGTTCGTATCGGCAACAGCGCGTCAATGACCAATGTAAGACTGGCGTAAGC 360
Qy 121 GilyLeuValLysGlyGlyGluAsPProGlnLysValAlaAlaAsPQinLeuValLysAP 140
Db 361 GGTTCGTCAAGAGGGGTGAAGATCCGCAAAAAGTGGCGCTGATCAATGTGTTAAGAT 420
Qy 141 GilyValAsPileuileThrileGlyGlyAsPAsPThraSnThraIaIleAsPLeu 160
Db 421 GGTCGTGATATTCGACACCATCGCGGGGATGATGATCAATACCGACAGCGGATTTG 480

Qy 161 AlaIaPheLeuAlaArgAsnAntTyGlyLeuThraValileGlyLeuProlyseThraVal 180
Db 481 GCAGATTCCTGGCCAGAAATATATACGAGCTGACCGCTATGTGTTCTTAAACCGCTC 540
Qy 181 AsPAnPAsPValPheProIlelyeGlnSerLeuGlyAlaIleTPThraAlaileGluGlnGly 200
Db 541 GATTAACGAGTATTTCCATCAAGCAATCACTAGGTGCTGAGTCTGCGCGCAAGCAGC 600
Qy 201 AlaArgTyPheMetAsnValAlaileGluAsnAsnAlaAsnProArgMetLeuileVal 220
Db 601 GCGCTTATTTCAATGAACGTGTGGCGCAAAAACAACGCCAACCCGATGCTGATCGTA 660
Qy 221 HisGlyValMetGlyArgAsnCyseGlyTyPLeuThraAlaIleThraAlaGlnGlyTyArg 240
Db 661 CACGAAGGAGAGGGCGGTACTGCGGCTGACCGCTGACCGCTGCAACCGCAGGAATATGCGC 720
Qy 241 LysLeuLeuAsPAsPValaGluTyPLeuProGluLeuGlyLeuThraArgGlySerTyGlu 260
Db 721 AAATTAATCTGAGCCGTCGAGAGTGTGCGGAATTTGGTTGACTCTGTAATCTTATGAA 780
Qy 261 ValHisAlaValPheValProGluMetAlaileAsPLeuGluAlaileLysArgLeu 280
Db 781 GTGCACGGCGGTATTCGTTCCGGAATATGCGATCGACTGGAAGCCGAAGCCAGCGCTG 840
Qy 281 ArgGlyValMetAsPLeuValAsPcyValAsnIlePheValSerGlyValaileGlyVal 300
Db 841 CGCGAAGTGAATGACAAAGCTGATTCGCTCAACCTTCGTTCCGAAGGTGCGCGCTC 900
Qy 301 GilyAlaileValAlaileGluMetGlnAlaileGlyGlnGluValProArgAsPAlaPheGly 320
Db 901 GAGCTATCTGCGCGAAATCGAGCCCAAGGCGCAAGAGTGCAGCGGATGCTTCGCGC 960
Qy 321 HisIleLysLeuAsPAlaValaAsnProGlyTyTrPheGlyGluGlnPheAlaGlnMet 340
Db 961 CACATCAAACTGAGATGCGGTCAACCTGTGTAATGTTCCGGGAGCAATTCGCGCAATG 1020
Qy 341 IleGlyAlaGlyTyThraLeuValGlnLysSerGlyTyPheAlaArgAlaSerAlaSer 360
Db 1021 ATAGCGCGGAAAAACCTGCTGACMAAAATCGGGATCTTCGCGCGGCTTCCTCC 1080
Qy 361 AsnValaAsPAsPMetArgLeuileLysSerCyAlaAsPLeuAlaValaileGlyCyAlaPhe 380
Db 1081 AACGTTGACGACATGCGTTGATCAATCTGCGCGGCTTGGCGGTGAGTCCGCTTC 1140
Qy 381 ArgArgGlySerGlyValileGlyHisAsPQinAsPAsnGlyAsnValLeuArgAlaile 400
Db 1141 CGCCGCAAGTCTGGGTATCGGTACAGCAAGCAACCGCAAGCTGTGTGTCGATC 1200
Qy 401 GIupheProarGilelyeGlylybProPheasniLeasPThAsPThPheasner 420
Db 1201 GAGTTTCGGCGATCAAGGGCGGCAACCGTTCAATATCGACACCGACTGTTCAATAGC 1260
Qy 421 MetLeuSerGluileGlyGlnProlyseGlylybValGluValSerHis 437
Db 1261 ATGTTGAGGAATCGCGCCAGCTTAAGCGGTAAAGTCGAAGTCAGCCAC 1311
RESULT 7
US-10-321-210-15
/ Sequence 15, Application US/10321210
/ Publication No. US20040115657A1
/ GENERAL INFORMATION:
/ APPLICANT: Kofifas, Matheos
/ APPLICANT: Odom, J. Martin
/ APPLICANT: Norton, Kelley C.
/ APPLICANT: Ye, Rick
/ TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
/ FILE REFERENCE: CL1619 US NA
/ CURRENT APPLICATION NUMBER: US/10/321,210
/ PRIOR FILING DATE: 2002-12-17
/ PRIOR APPLICATION NUMBER: US/09/934,901
/ PRIOR FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/229,906

; PRIOR FILING DATE: September 1, 2000
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 15
 ; LENGTH: 1311
 ; TYPE: DNA
 ; ORGANISM: METHYLOMONAS SP.
 ; US-10-321-210-15

Alignment Scores:

| Pred. No.: | 4,19e-253 | Length: | 1311 |
|------------------------|-----------|---------------|------|
| Score: | 2284.00 | Matches: | 437 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 18 | Gaps: | 0 |

US-10-701-200-6 (1-437) x US-10-321-210-15 (1-1311)

```

QY 1 AppValValThrTrpProTyrHisLeuThrAlaAspIleArgPheCysHisTrpPhe 20
DB 1 GATGGTGCACATGGCCCTATCATCTTAACGGCTGATTCGATTGTCATTGGTTTTT 60
QY 21 LeuAsnProAsnPhenylTrpHisLeuMetAsnLysProLysLysValAlaIleLeuThrAla 40
DB 61 CTTAACCTTAACTTCTTACACGCTCATGAACAACTTAAAGTTCCAACTACTGACAGCA 120
QY 41 GlyLysLeuAlaProCysLeuAsnSerAlaIleGlySerLeuIleGluArgTyrThrGlu 60
DB 121 GCGCGCTTGGCCCTTGTTGTAATCCGCAATCGAGTTGTTGATTCAGAACGTTATACGAA 180
QY 61 IleAspProSerIleGluIleIleCysTyrArgGlyGlyTyrLysGlyLeuLeuLeuGly 80
DB 181 ATCGATCCAGCATGAATGATCAATTCCTATCCGGCGGTTATTAAGCCCTGGTGGGC 240
QY 81 AppSerTyrProValThrAlaGluValArgLysValGlyValLeuGluArgPheGly 100
DB 241 GATTCCTTATCCAGTAAACGCGCCGAAGTCGTAAGAGCGCGGTTCCTGCACGTTTGGC 300
QY 101 GlySerValIleGlyAsnSerArgValLysLeuThrAsnValLysAspCysValLysArg 120
DB 301 GGTCTGTGATGGGCAACGCGCGCTCAATGACCAATGTCAGAACATGCGTGAACCC 360
QY 121 GlyLeuValLysGlyGluLysAspProGluLysValAlaAlaAspGluLeuValLysAsp 140
DB 361 GGTGGTGTCAAGAGGGGTGAAGATCCGCAAAAAGTCCGGCTGATCAATTGGTTAGCAT 420
QY 141 GlyValAspIleLeuHisThrIleGlyGlyAspAspThrAsnThrAlaAlaAspLeu 160
DB 421 GGTGTGATATCTCTGCACACCATCGCGCGCATGATACCAATTCGCGCACAGCGGATTG 480
QY 161 AlaAlaPheLeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLysThrVal 180
DB 481 GGAGCATTCCTGGCCGGAATAATTAACGACGATGATGGTATTACCTAAACCGTC 540
QY 181 AspAsnAspValPheProIleLysGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGly 200
DB 541 GATTAACGACGATATTCCTCGATCAAGCAATCACTAGTGTCTTGACCTCCCGCAGCAAGC 600
QY 201 AlaArgTyrPheMetAsnValAlaAlaGluAsnAsnAlaAsnProArgMetLeuIleVal 220
DB 601 GCGCGTATTTATTCATGAACGTGGTGGCGCAAAACCAAGCCCAACGATGCTGATCGTA 660
QY 221 HisGluValMetGlyArgAsnCysGlyTyrPheLeuThrAlaAlaThrAlaGlnGluTyrArg 240
DB 661 CACGAAGTATGGCGCGTAACTGGCGCTGATGACCGCTGCAACCCCGCAGGAATATCC 720
QY 241 LysLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArgGlnSerTyrGlu 260
DB 721 AATTAATCTGACCGCTCCGAGTGGTGGCGGAATTTGGGTTTGACTGTGAATTTATGA 780
QY 261 ValHisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeu 280

```

```

DB 781 GTGCACGCGCTATTGCTTCGGAATATGCGCATGACCTGAAAGCCGAACCAAGCCCTG 840
QY 281 ArgGluValMetAspLysValAspCysValAsnIlePheValSerGluValAlaGlyVal 300
DB 841 CGCGAAGTATGGACAAAGTCGATTCGCTCAACATCTTCTCCGAAGTGGCGCGCTG 900
QY 301 GluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGly 320
DB 901 GAAGCATGTCGCGGAATATGCAAGGCCAAGAGTCCGCGCATGCTGCTCGC 960
QY 321 HisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGlnGlnPheAlaGlnMet 340
DB 961 CACATTAACCTGATGCGGTCAACCTGTAATATGTTGGGAGCAATTCGCGCAATG 1020
QY 341 IleGlyAlaGluLysThrLeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSer 360
DB 1021 ATAGCGCGCGAATAAACTGTTACAAATATCGGATCTTGGCCGCTCTGCTTCC 1080
QY 361 AsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPhe 380
DB 1081 AACGTTGACGACATGCGTTGATCAATCTGCGCGCATGCGGTGCGATGCGCGTTC 1140
QY 381 ArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsnValLeuArgAlaIle 400
DB 1141 CCGCGGAGTCTGGCTGTATCGGTCAAGCAAGACCAACGCAACGTTGTGCGTGATC 1200
QY 401 GluPheProArgIleLysGlyGlyLysProPheAsnIleAspThrArgPheAsnSer 420
DB 1201 GATTTCCGCGCATCAAGGCGCGCAACCGTTCAATTCGACACCGATCGTTCAATAGC 1260
QY 421 MetLeuSerGluIleGlyGlnProLysGlyGlyLysValGluValSerHis 437
DB 1261 ATGTTGAGCGGAATCGCGCAGCCTAAAGCGGTAAAGTGAAGTCAGGCAC 1311

RESULT 8
US-10-369-493-38035
; Sequence 38035, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38035
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; US-10-369-493-38035

Alignment Scores:
Pred. No.: 4,51e-66 Length: 552
Score: 659.00 Matches: 125
Percent Similarity: 82.61% Conservative: 27
Best Local Similarity: 67.93% Mismatches: 32
Query Match: 28.85% Indels: 0
DB: 17 Gaps: 0

US-10-701-200-6 (1-437) x US-10-369-493-38035 (1-552)
QY 33 LysLeuValAlaIleLeuThrAlaGlyGlyLysLeuAlaProCysLeuAsnSerAlaIleGly 52
DB 1 CAGAAAGTGCAGATGCTGACCGCGGTGCTCGCGCCCTGCTCTTCCGCGCGGT 60
QY 53 SerLeuIleGluArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGly 72

```

```

Db      61  |||||.....||| |||.....|||
GCGCTGATCGAGCGGTATAGCAGCATCGCGCTGAAATCGACATATCGCTTACCGCTCC 120
Qy      73  G|YrYrYsgYlYleuLeuLgYlYAspSerTYrProValThrAlaGluValArgYlYs 92
Db      121  GGTATATACGGGGGTCTCTTGGGCGAGGATATCGAGATCACCAAGATATGGCGAAAG 180
Qy      93  AlAgYlValLgUlnArgPheGlyYserValIleGlyAsnSerArgValYlYLeuThr 112
Db      181  GCGCATCTACTGCAATCTTACCGCGCTCGCGATCGGCAACAGCGCGTCAAGCTGAC 240
Qy      113  AsnValYlYAspCYsValYlYAspArgYlYLeuValYlYsGluGlyAspProGluYlYVal 132
Db      241  ATGCGCGCGATTCGCCAGAGCGCGCTCGTCAAGAGAGGAGCAATCGTTGCGCGTC 300
Qy      133  AlAlaAspGlnLeuValYlYAspGlyValAspIleLeuHlSerHlIleGlyYlYAspAsp 152
Db      301  GCGCGTGAAGCGCTCGCGCGTACGCGCATCACTTCCTCAATCACTATCGGTGCGACAC 360
Qy      153  ThrAsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTYrGlyLeuThr 172
Db      361  ACCAAGACCGCGCGCGCGCTTGCCTCTCTCGCGCGCGCAACGATATGACCTTCAAC 420
Qy      173  ValIleGlyLeuProYlYsThrValAspAsnAspValPheProIleYlYsGlnSerLeuGly 192
Db      421  GTGCGGCGCTGCGAGAGCGGTGATATGACGTGCGCGCATCAAGAGTGGCTGCGC 480
Qy      193  AlATrThrAlaAlaGluGlnGlyAlaArgTYrPheMetAsnValAlaGluAsnAsn 212
Db      481  GCGTGAAGCGCGCGGAGTGGTCTGCTCTTTCGACAAATGTGACAGACAGACAGAC 540
Qy      213  AlAsnProArg 216
Db      541  GCGGACCGGAAA 552

```

RESULT 9

```

US-10-369-493-31795
; Sequence 31795, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31795
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Thermobifida fusca
US-10-369-493-31795

```

Alignment Scores:

```

Pred. No.: 5,696-25 Length: 1050
Score: 306.50 Matches: 104
Percent Similarity: 42.45% Conservative: 63
Best Local Similarity: 26.46% Mismatches: 145
Query Match: 13.42% Indels: 81
Db: 17 Gaps: 14

```

US-10-701-200-6 (1-437) x US-10-369-493-31795 (1-1050)

```

Qy      23  PheAsnPheTYrThrLeuMetAsnYlYsProYlYsValAlaIleLeuThrAlaGlyGly 42
Db      10  TTCAAGAGACTACCGCAATG-----CGAGTGGGGTACTACTACCGGGGGTGGC 57

```

```

Qy      43  LeuAlaProCYsLeuAsnSerAlaIleGlySerLeuIleGluArgTYrThrGluIleAsp 62
Db      58  GACTGCCGGGCGCTGAAGCGGTGATCGGTGCGGGTCCCTTGAAGGCAATCAAG--GAA 114
Qy      63  ProSerIleGluIleIleCYsTYrArgGlyGlyTYrYlYsGlyLeuLeuLgYlYAspSer 82
Db      115  TACGATACAGAGTTCTGTCGATTCGTGACGGGTGGCGGCGCCCTGGAAAGCGACACC 174
Qy      83  TYrProValThrAlaGluValArgYlYsAlaGlyValIleGluArgPheGlyYlYser 102
Db      175  ATGCGCGTGAATCGAGCGGTGCGC-----GGCATCTCCCGCGC--GGCGGTACC 225
Qy      103  ValIleGlyAsnSerArgValYlYsLeuThrAsnValYlYsAspCYsValYlYsArgYlYLeu 122
Db      226  ATCTCGGCTCTCGCGGACCAACCTCTGAAAGATCGAGGGCGGTGACGCG--G-- 279
Qy      123  ValYlYsGluGlyGluAspProGluYlYsValAlaAlaAspGlnLeuValYlYsAspGlyVal 142
Db      280  -----GTCAAGACAAATGCGCGCGCTCGGGGTG 309
Qy      143  AspIleLeuHlSerHlIleGlyYlYsAspThrAsnThrAlaAlaAspLeuAlaAla 162
Db      310  GACGCGCTCGTCCGAATCGCGGTGAGGACACCTCGCGCTGCGCGCACGCTC----- 363
Qy      163  PheLeuAlaArgAsnAsnTYrGlyLeuThrValIleGlyLeuProYlYsThrValAspAsn 182
Db      364  -----CACGACACAGGTGTCAACGTGTGGTGTGCGCCCAAGACATCGAACAC 411
Qy      183  AspValPheProIleYlYsGlnSerLeuGlyAlaTrpThrAlaAlaGluGlnGlyAlaArg 202
Db      412  GACCTCAAGCGCACCGACTACCTTCGCTTCCAGCATCTGGGTGATCAATCGGACCGAG 471
Qy      203  TYrPheMetAsnValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHlGlu 222
Db      472  GCCATCGACCGCTGACACACCGCGCGAGTCGACACACGCTGCGCTGCTGCTC--GAG 528
Qy      223  ValMetGlyArgAsnCYsGlyTYrLeuThr-----AlaMetThrAlaGln 237
Db      529  GTCATGGGCGCGCGCGGTGATTCGCTGACGCGGCGCATGCGCGCGCGCGCC-- 585
Qy      238  GluTYrArgYlYsLeuLeuAspArgAlaGluTrpLeuProGluLeuGlyYlYLeuThrArgGlu 257
Db      585  ----- 585
Qy      258  SerTYrGluValHlAlaValAlaPheValProGluMetAlaIleAspLeuGluAlaGluAla 277
Db      586  -----AACGTCACTCATCCCGAGCGCCCATTCGACATCGACGAGGTGCTC 633
Qy      278  YlYsArgLeuArgGluValMetAspYlYsValAspCYsValAsnIlePheValSerGluGly 297
Db      634  GCTTACATCGAAGCGCGTTCAAGACCACTACGCGCGCGATCATGTGATCGCCAGAGGC 693
Qy      298  AlAgYlValGluAlaIleValAlaGluMetGlnAlaYlYsGlyGlnGluValProArgAsp 317
Db      694  GCCACCGGAG-----GAGGCGCAGCTCACCTCGGCGGACGCGGAGACGCGAC 741
Qy      318  AlAphGlyHlSerIleYlYsLeuAspAlaValAsnProGlyYlYsTrpPheGlyGluGlnPhe 337
Db      742  TCCTTCGCGCAGTCCGCGCTCGCGCGCATC-----GCCAGCGCGCTC 783
Qy      338  AlAglnMetIleGlyAlaGluYlYsThrLeuValGlnYlYsSer-----GlyTYrPhe 354
Db      784  GCCGAGAGATTCGAGCGCGGACCGGACGCGGACGCGCGGTGCTGCTGCGGACCGTG 843
Qy      355  AlAtrGluAsnSerAlaSerAsnValAspAspMetArgLeuIleYlYsSerCYsAlaAspLeu 374
Db      844  CAGCGCGCGGTAGCGCTTCGCTTCGACCGGGTGTGGCCACCGGCTGGGCTGGC 903
Qy      375  AlAvalGluCYsAlaPheArgArgGluSerGlyValIleGlyHlAspGluAspAngly 394
Db      904  GCGATCAACGGGGT-----CACGACAAAGACCTTGGC 936
Qy      395  AsnVal-----LeuArgAlaIleGluPheProArgIle 405

```


APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 6060
LENGTH: 1026
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1026)
US-10-156-761-6060

Alignment Scores:
Pred. No.: 1,526-23 Length: 1026
Score: 294.00 Matches: 101
Percent Similarity: 43.83% Conservative: 66
Best Local Similarity: 26.51% Mismatches: 140
Query Match: 12.87% Indels: 74
DB: 15 Gaps: 17

US-10-701-200-6 (1-437) x US-10-156-761-6060 (1-1026)

QY 34 LysValAlaIleLeuThrAlaGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53
DB 4 CGGGTCGAGTACTACCGGAGCGCGAGCTGCCCGGGCTAACGCCGTCATCCGGGCG 63
QY 54 LeuIleGluArgTyrThrGluIleAspProSerIleGluIleCysTyrArgGly 73
DB 64 CTCGTCCTCGTAAGGGGTGCAG--GAGTACGGCTACGACTTGTGGCTCCGGAGCGC 120
QY 74 TyrLysGlyLeuLeuGluGlyAspSerTyrProValThrAlaGluValArgLysAla 93
DB 121 TGGCGGGGACCACTCGAAGCGAGCGC-----GTACGGCTCGACATCCCGGGTGGC 174
QY 94 GlyValIleGluGlnArgPheGlyGlySerValIleGlyAsnSerArgValLysLeuThrAsn 113
DB 175 GGCATCTCGCCCCGCG--GGCGGACGATCTCGGCTCTCGCGACCAACCCCTCAAG 231
QY 114 ValLysAspCysValLysArgGlyLeuValLysGluGlyGluAspProGlnLysValAla 133
DB 232 CTCGACGACGCGATCCGCGG-----ATCAAGAG----- 261
QY 134 AlaAspGlnLeuValLysAspGlyValAspIleLeuHisThrIleGlyGlyAspAspThr 153
DB 262 -----AACCTCGCCCAAGCAGAGAGGTGCGCGCTCATCGGATCGCGGCGAGGACAG 315
QY 154 AsnThrAlaAlaAlaAspLeuAlaAlaPheLeuAlaAlaArgAsnAsnTyrGlyLeuThrVal 173
DB 316 CTCGGCGCTCGCGCAGCGCTCAC-----GACGAGTACGCGCTGCGCGCGC 360
QY 174 IleGlyLeuProLysThrValAlaAsnAspValPheProIleLysGlnSerLeuGlyAla 193
DB 361 GTCGGCGCTCCGAAGACCATCGACACGACTGTGCGGACCGACATCACGTTCCGCTTC 420
QY 194 TrpThrAlaAlaGluGlnGlyAlaArgTyrPhe-----MetAsnValValAlaGluAsn 211
DB 421 GACACGGCGCGTGGGATCGGACCGAGCGCATCGACCGCTCCACACCGCCGAGTCC 480
QY 212 AsnAlaAsnProArgMetLeuIleValHisGluValMetGlyArgAsnCysGlyTyrLeu 231
DB 481 CACATGCGC-----GTCTCTGCTCGGAGGTGATGGCGCGGACCGCGGCTGATC 531
QY 232 ThrAlaAlaThrAlaGlnGluTyrArgLysLeuLeuAspArgAlaGluTyrLeuProGlu 251
DB 532 GCCATCCACTCG----- 543

QY 252 LeuGlyLeuThrArgGluSerTyrGluValHisAlaValPheValProGluMetAlaIle 271
DB 544 ---GGCTGGCGCGCGCGC-----AACGTCACTTCATCCCGAGGAGCGCTTC 591
QY 272 AspLeuGluAlaGluAlaLysArgLeuArgGluValMetAspLysValAspCysValAsn 291
DB 592 GACGTGACCAAGGTGTGCGCTATGTACCTCGCGGTTCAAAGGCTCGTACCGCGCATC 651
QY 292 IlePheValSerGluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAlaLysGly 311
DB 652 GTGTGTGTCGCGGAGGGCGC-----ATGCCAAGACAGCG 687
QY 312 GlnGluValProArg-----AspAlaPheGlyHisIleLysLeuAspAlaVal 327
DB 688 CAGATGGTCTCAAGAGACAGATCCCTCGACTTCCTCGGACAGTCCGCTTCGCGGGTC 747
QY 328 AsnProGlyLysTrpPheGlyGluGlnPheAlaGluMetIleGlyAlaGlu---LysThr 346
DB 748 -----GGCGAGTGGCTGGCCCAAGAGATCGAAGAGCCACCGGACAGAGCCCGCAC 801
QY 347 LeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValAspAspMetArg 366
DB 802 ACGGTG-----CTCGGCGCATGTGACGCGCGGACGCGGAGCGCGTTCGACCGTGG 855
QY 367 LeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArgGluSerGlyVal 386
DB 856 CTCGCCACCCGCTTCGGCTCGACGCGCATCGAGGCGCTC----- 894
QY 387 IleGlyHisAspGluAspAsnGlyAsnVal-----LeuArgAlaIleGluPheProArg 404
DB 895 -----CGCGACGGCGACTTGGCGCAAGATGGTCCGCTCGCGGACGACATCGTCCGT 948
QY 405 Ile 405
DB 949 GTC 951

RESULT 12
US-10-425-114-9204
Sequence 9204, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 9204
LENGTH: 1570
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700837684_FLI
US-10-425-114-9204

Alignment Scores:
Pred. No.: 3,716-23 Length: 1570
Score: 293.00 Matches: 113
Percent Similarity: 41.23% Conservative: 61
Best Local Similarity: 26.78% Mismatches: 174
Query Match: 12.83% Indels: 74
DB: 17 Gaps: 15

US-10-701-200-6 (1-437) x US-10-425-114-9204 (1-1570)

QY 34 LysValAlaIleLeuThrAlaGlyLeuAlaProCysLeuAsnSerAlaIleGlySer 53


```

Db 224 CAGGCTGCTATTGTTACTTGTGGGGTCTGTCTCGGGCTCAACACTGCTCATTTAGGGA 283
Oy 54 Leuilegluarglythrthrgluileaprosertleleuilelecystrhrarglygly 73
Db 284 TTAGTGTGTGGCTTACACATATGTTAGGGGGAAGAAAGTTCTTGGAATCAATGAGAGA 343
Oy 74 Trrlyeglyleuleuleuileuileapserlyrprovalthrslagluvalarglylyvala 93
Db 344 TATAGGGGTTTCTATGCTCGCATATACATCACTTTAAACCCCT-----AAAAGTGTG 394
Oy 94 Glyvalleuileuileuileuileuileuileuileuileuileuileuileuileuileu 113
Db 395 AATGATATATCATAGAGGCTGGGGAAGTCTCTCGAATCATACGA----- 439
Oy 114 Vallyabpsercyvallyabpserglyleuvallybgluglygluabpseroglinlyvala 133
Db 440 -----GGTGCACATGACACCAAAAGATA--- 463
Oy 134 Alaabpserleuvallyabpserglyvalaaprlieleuilethrileglylyabpserthr 153
Db 464 GTTGACAGATTCAAGATCGGGGAATCAATCAAGTTATATATGAGAGAGATGGAAGT 523
Oy 154 Aenthrallalaalabpserleuallalaarpeleuallargabpserlyrlyleuthrval 173
Db 524 CAGAAAGGTGCATCTGCATTTTGTGAGGAAGTCAGAAAGCTGTCTCAAGTTTCAGTT 583
Oy 174 Ileglyleuprolyethrvalabpserapvalphproilleuileginsertleuileuile 193
Db 584 GTAGGATATCCCAAAACATAGATATGATATGATTCAGTTATGATAGTCTTTGGCTTT 643
Oy 194 Trrthrallalaagluugluvalaarglyrphemeleuvalvalaagluabpserleu 213
Db 644 GACATCGCTGTTGAGGAGCTCAACAGCTATTAATGACGACATGTTGAGCTGAAGT 703
Oy 214 Aenproabpserleuilevalhilegluvalmetglyabpsercyglyrtrpleuthrvala 233
Db 704 GTAGAAATGAGCATAGTGTGTGTCAGTTGATGAGTGAAGAACGCGATTTATTTGCA--- 760
Oy 234 Alathrralagluuglyrarglyleuileuabpseralagluirpleuoprogluileuile 253
Db 761 -----ATGTAT 766
Oy 254 Leuthrarglyserlyrgluvalhilevalalevalaleuileuileuileuileuileu 273
Db 767 GCTACTCTTGCGAGTGCAGACGTGATGTGTGCTTATTCAGAGTCAACCTTTTACTTT 826
Oy 274 Gluvalagluale-----lyabpserleuarglyluvalmetabpserly 286
Db 827 GAAGGTCCTGTGACTCTATGAAATATATAGAGAAAGACTAAAGAAATGGGCACATG 886
Oy 287 Valabpsercyvalaasnhleuileuileuileuileuileuileuileuileuileuileu 306
Db 887 GTT-----ATTGTATGTCTGAAGGACGACAGAACCTTTCTCTGAGAGT 934
Oy 307 Metglnaleuileuileuileuileuileuileuileuileuileuileuileuileuileu 326
Db 935 GTGCACTCATGACGAAACAA-----GATGCTTCTGGAACCAAGCTTTTCAAGAT 985
Oy 327 Valaapseroglylyrtrpserglyuileuileuileuileuileuileuileuileuileu 346
Db 986 GTT-----GGCCTATGATATCCCAAAAGATTAGGATCATTTTCTGACAGAGACA 1039
Oy 347 Leu-----valglnyabpserglyrphleuileuileuileuileuileuileuileu 361
Db 1040 CTGCCCATATCTCTCAATATACATAGATCCAACTATATATATCGACCTATTTCCAAACCAT 1099
Oy 362 Valaapsermetabpserleuileuileuileuileuileuileuileuileuileuileu 381
Db 1100 GGCTGTGACAGATGATGACGACATCTTCTGCAAAAGTGCAGTTGATGAGAGAGAG--- 1156
Oy 382 Arggluserglyvalileglynhlabpsergluabp-----Aenplyabpvalleuargla 399

```

```

Db 1157 -----GACGTTTACATGCTGCTATACAGTGCATTTGCAATGAGAGA---CAAACTTAT 1207
Oy 400 Ilegluabpserarglyleuileuileuileuileuileuileuileuileuileuileu 417
Db 1208 ATACCTTATATAGATATCATGAGAGACGAAACAGTATGATATATGATGAAATGTGG 1267
Oy 418 Pheabpsermetleuileuileuileuileuileuileuileuileuileuileuileuileu 432
Db 1268 GCTAGG-----CTTTTATCTTCAACAAATCAACCCAGCTTTTGATGATCCAAAGGTGACAAT 1324
Oy 433 Valglu 434
Db 1325 GAAGAG 1330

RESULT 13
US-10-425-114-6293
/ Sequence 6293, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 6293
/ LENGTH: 1765
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700559917_FLI
US-10-425-114-6293

Alignment Scores:
Pred. No.: 4,41e-23 Length: 1765
Score: 293.00 Matches: 113
Percent Similarity: 41.23% Conservative: 61
Bee Local Similarity: 26.78% Mismatches: 174
Query Match: 12.83% Indels: 74
DB: 17 Gaps: 15

US-10-701-200-6 (1-437) x US-10-425-114-6293 (1-1765)
Oy 34 Lyvalalaileuileuileuileuileuileuileuileuileuileuileuileuileuileu 53
Db 371 CAGGCTGCTATTGTTACTTGTGGGGTCTGTCTCGGGCTCAACACTGCTCATTTAGGGA 430
Oy 54 Leuilegluarglythrthrgluileaprosertleleuilelecystrhrarglygly 73
Db 431 TTAGTGTGTGGCTTACACATATGTTAGGGGGAAGAAAGTTCTTGGAATCAATGAGAGA 490
Oy 74 Trrlyeglyleuleuleuileuileuileuileuileuileuileuileuileuileuileu 93
Db 491 TATAGGGGTTTCTATGCTCGCATATACATCACTTTAAACCCCT-----AAAAGTGTG 541
Oy 94 Glyvalleuileuileuileuileuileuileuileuileuileuileuileuileuileu 113
Db 542 AATGATATATCATAGAGGCTGGGGAAGTCTCTCGAATCATACGA----- 586
Oy 114 Vallyabpsercyvallyabpserglyleuvallybgluglygluabpseroglinlyvala 133
Db 587 -----GGTGCACATGACACCAAAAGATA--- 610
Oy 134 Alaabpserleuvallyabpserglyvalaaprlieleuilethrileglylyabpserthr 153
Db 611 GTTGACAGATTCAAGATCGGGGAAGTCTCTCGAATCATACGA----- 670

```


US-10-701-200-6 (1-437) x US-10-424-599-141147 (1-4413)

```

OY      274  |||u|l|a|g|u|a|e-----|y|a|r|e|u|a|g|l|u|a|l|e|a|p|y|e 266
Db      1166 GAAAGCTCTGTGGACCTGTATGAATATATAGAGAAAGACTTAAAGAAATGGGCACATG 1225
OY      287  valaPcYsValaEni|ePheValSer|u|l|y|a|g|l|u|a|l|e|a|l|a|g|u 306
Db      1226 GTT-----ATTGTTATTTGCGTGAAGACAGACAGAACTGTTTCTGAGAGT 1273
OY      307 Met|l|n|a|l|a|y|g|l|n|g|l|u|a|l|P|o|r|a|r|g|a|P|a|P|he|g|y|h|i|s|i|e|l|y|e|u|a|P|a|a 326
Db      1274 GTGAGTCCTCAGACCAACAA-----GATGCTTCTGGAAACAAGCTTTTCAAGAT 1324
OY      327 ValaEnP|o|r|g|l|y|s|T|P|P|he|g|l|y|g|l|n|g|l|n|P|he|a|l|a|g|i|n|e|t|i|e|g|y|a|l|a|g|l|u|e|t|h|r 346
Db      1325 GTT-----GGCTATAGGATATTC|C|C|A|A|A|A|G|T|T|A|G|G|A|G|A|C|A|T|T|T|G|C|T|G|C|A|C|A|G|A|C|A 1378
OY      347 Leu-----ValG|n|L|y|s|e|r|g|l|y|T|r|P|he|a|l|a|r|g|a|l|a|s|e|r|a|n 361
Db      1379 CTGCCCATTACTCTCAAAATATATAGATTCACACCTATATGATCCAGAGTATTCACAGCAAT 1438
OY      362 ValaP|a|P|e|c|a|r|g|l|e|l|y|e|s|e|C|y|e|a|l|a|P|e|u|a|v|a|l|g|u|C|y|a|l|a|P|e|a|r|g 381
Db      1439 GCCCTCTGACACAGTGTACTGACACACTTCTTGCTCAAAAGTCAGCTTCATGAGCAATG--- 1495
OY      382 Arg|l|u|s|e|r|g|l|l|e|g|l|y|h|i|a|P|o|l|u|a|r|-----A|e|n|g|y|a|n|V|a|l|e|u|a|A|r|g|a| 399
Db      1496 -----GAGGTATACACTGCTGTATACAAAGTGGACTTGTGAATGGAGAA---CAAACTTAT 1546
OY      400 l|e|g|l|u|P|e|P|o|r|a|r|g|l|e|l|y|g|l|y|e|P|o|r|P|e|a|n|i|e|a|P|Th|a|r|P|-----T|r|p 417
Db      1547 ATACCTTCTTATATATATCATCAGAGACAGAAACACAGTATGATGATTAATGATGCG 1606
OY      418 P|he|a|n|S|e|r|v|e|t|e|u|s|e|r|g|l|l|e|g|l|n|P|o|-----L|y|e|g|l|y|L|y|s 432
Db      1607 GCTAGG---CTTTATCTTTCACAAATCAACCCAGCTTTTGGATGCCAAGGTGACAAAT 1663
OY      433 Val|g|u 434
Db      1664 GAAGAG 1669

RESULT 15
US-10-424-599-141147
US-10-424-599-141147, Application US/10424599
US-10-424-599-141147, Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalick David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 141147
LENGTH: 4413
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_98468C.1
US-10-424-599-141147

Alignment Scores:
Pred. No.: 1,69e-22 Length: 4413
Score: 293.00 Matches: 113
Percent Similarity: 41.23% Conservative: 61
Best Local Similarity: 26.78% Mismatches: 174
Query Match: 12.83% Indels: 74
DB: 17 Gaps: 15
```

```

Qy 382 ArgGluSerGlyValIleGlyHisAspGluAsp-----AsnGlyAsnValLeuArgAla 399
Db 3887 -----GCAGGTTACACTGGCTATACAAGTGGACTTGTGAATGAAGA---CAAATTAT 3937
Qy 400 IleGluPheProArgIleLysGlyLysProPheAsnIleAspThrAsp-----Trp 417
Db 3938 ATACCTTCTATAGAAATCACTGAGAGACAGAACACAGTAGTGATATGAATGATG 3997
Qy 418 PheAsnSerMetLeuSerGluIleGlyGlnPro-----LysGlyGlyLys 432
Db 3998 GCTAGG---CTTTATCTTCACAACAATCAACCCAGCTTTTGTGATGCCAAGGTGACAAT 4054
Qy 433 ValGlu 434
Db 4055 GAAGAG 4060

```

Search completed: March 1, 2005, 21:18:32
 udb time : 4680 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2005, 14:11:02 / Search time 4373 Seconds
(without alignments)
3803.816 Million cell updates/sec

Title: US-10-701-200-6

Perfect score: 2284
Sequence: 1 DVATPMTYLTADIRCFHWF.....FNSMLSEIQPKGKXVESH 437

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q/cgn2.1/USPTO_spool/US10701200/runat.25022005.102832.10319/app.query.fasta.1.583
-DB=EST -QMT=fastlap -SUFFIX=trc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-OCCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10701200 @CNC 1.1 3437 @runat.25022005.102832.10319 -NCPU=6 -ICPU=3
-NO MAP -LARGEOUTER -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | ID | Description |
|------------|-------|--------------|-----------------|-------------|
| 1 | 1341 | 58.7 | 1177 2 BE636682 | rockefeller |
| 2 | 736 | 32.2 | 881 6 CD379187 | PTMM04356 |
| 3 | 712 | 31.2 | 873 6 CD383878 | PTMM09050 |
| 4 | 671 | 29.4 | 743 4 BM321461 | rockefeller |
| 5 | 665.5 | 29.1 | 896 6 CD379057 | PTMM04226 |
| 6 | 629 | 27.5 | 878 6 CD379942 | PTMM00103 |
| 7 | 597.5 | 26.2 | 878 6 CD383585 | PTMM08757 |
| 8 | 588 | 25.7 | 904 6 CD382389 | PTMM07559 |
| 9 | 565.5 | 24.8 | 932 6 CD378198 | PTMM03367 |

| | | | | |
|----|-------|------|------------------|-----------|
| 10 | 564 | 24.7 | 746 6 CD384626 | CD384626 |
| 11 | 543.5 | 23.8 | 757 6 CD375650 | PTMM09799 |
| 12 | 516.5 | 22.6 | 639 6 CD380358 | CD375650 |
| 13 | 513.5 | 22.5 | 916 6 CD379504 | PTMM05247 |
| 14 | 506 | 22.2 | 566 6 CD379504 | PTMM04673 |
| 15 | 485.5 | 21.3 | 677 6 CD378314 | CD378314 |
| 16 | 483.5 | 21.2 | 677 6 CD375439 | PTMM03483 |
| 17 | 475.5 | 20.8 | 604 6 CD379552 | CD375439 |
| 18 | 425 | 18.6 | 316 4 BM321460 | PTMM00602 |
| 19 | 394.5 | 17.3 | 575 6 CD380700 | CD379552 |
| 20 | 384.5 | 17.3 | 671 6 CD384340 | PTMM05869 |
| 21 | 381.5 | 17.3 | 1797 6 CD380085 | CD380700 |
| 22 | 281.5 | 12.2 | 1629 3 CD380405M | BM321460 |
| 23 | 268.5 | 11.8 | 303 4 BI307578 | CD384340 |
| 24 | 264.5 | 11.6 | 1734 3 CNS0A094 | CD380085 |
| 25 | 261.5 | 11.4 | 1596 9 CL957442 | CD380085 |
| 26 | 254 | 11.1 | 1388 3 CNS0A261 | CD380085 |
| 27 | 253 | 11.1 | 1737 3 CNS0A261 | CD380085 |
| 28 | 251 | 11.0 | 990 9 CG155078 | CD380085 |
| 29 | 245.5 | 10.7 | 882 9 CG171162 | CD380085 |
| 30 | 244 | 10.7 | 982 7 CK278709 | CD380085 |
| 31 | 241 | 10.6 | 1036 9 CNS07402 | CD380085 |
| 32 | 235 | 10.3 | 701 6 CA208890 | CD380085 |
| 33 | 233 | 10.2 | 920 7 CV211012 | CD380085 |
| 34 | 231.5 | 10.1 | 764 4 BG365459 | CD380085 |
| 35 | 231.5 | 10.1 | 841 9 CNS0752F | CD380085 |
| 36 | 231.5 | 10.1 | 1867 3 CNS0A305 | CD380085 |
| 37 | 230.5 | 10.1 | 628 6 CA157062 | CD380085 |
| 38 | 230.5 | 10.1 | 1661 3 CNS0A482 | CD380085 |
| 39 | 229 | 10.0 | 921 7 CV215682 | CD380085 |
| 40 | 228 | 10.0 | 1743 9 CL959679 | CD380085 |
| 41 | 227.5 | 10.0 | 802 6 CD849303 | CD380085 |
| 42 | 226.5 | 9.9 | 1430 3 CNS0A113 | CD380085 |
| 43 | 226 | 9.9 | 896 9 CCG59323 | CD380085 |
| 44 | 226 | 9.9 | 897 7 CV209353 | CD380085 |
| 45 | 226 | 9.9 | 900 7 CV219943 | CD380085 |

ALIGNMENTS

RESULT 1
LOCUS BE636682 1177 bp mRNA linear EST 03-JAN-2002
DEFINITION rockefeller.0.370 Maestigameoba balamuthi lambda ZAP II library
Maestigameoba balamuthi cDNA similar to pyrophosphate--fructose
6-phosphate 1-phosphotransferase, mRNA sequence.

ACCESSION BE636682 GI:18055199
VERSION BE636682
KEYWORDS EST.
SOURCE Maestigameoba balamuthi
ORGANISM Maestigameoba balamuthi

REFERENCE

1 (bases 1 to 1177)
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durrille,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Maestigameoba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

TITLE

On Aug 25, 2000 this sequence version replaced gi:9919793.

JOURNAL

Medline
PubMed
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 1177 Std Error: 0.00
POLYA=No.

FEATURES

source 1.1177
/organism="Maestigameoba balamuthi"
/mol_type="mRNA"

```

ORIGIN
/strain="ATCC 30984"
/db xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/notes="Syn: Phreatamoeba balamuthi"

```

Alignment Scores:
Pred No.:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 2.82e-152 | Length: | 1177 |
| Score: | 1341.00 | Matches: | 252 |
| Percent Similarity: | 81.28% | Conservative: | 52 |
| Best Local Similarity: | 67.38% | Mismatches: | 70 |
| Query Match: | 58.71% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

| | | | |
|----|-----|---|-----|
| QY | 3 | LYELVYValAlaIleLeuThrAlaGlyGlyLeuAlaProCysLeuAsnSerAlaIleGly | 52 |
| Db | 54 | AMACRGTGGCTGCTCACTCCGAGGCGCTCGCCCTGCTGTGCGGCGGTGGC | 113 |
| QY | 53 | SerLeuIleGluNArgTyrThrGluIleAspProSerIleGluIleIleCysTyrArgGly | 72 |
| Db | 114 | GGCCTCATCGAGGCTACACCGAGATGTCGCCGAGACCAACATCATCTTACCTCAAC | 173 |
| QY | 73 | GlyTyrIleGlyLeuLeuLeuGluIleAspSerTyrProValThrAlaGluValArgLysLys | 92 |
| Db | 174 | GCGTCAAGAGCGCTGCTGCTGGAGGAAGTGTCGTGACCCCGGCATAGGCTCAG | 233 |
| QY | 93 | AlaGlyValIleuGluNArgPheGlyGlySerValIleGlyAsnSerArgValIleuLeuThr | 112 |
| Db | 234 | GCGCAAGTCTCTGCACACCGTCGGCGGACGTGCACTGGCAACTCGCGTCAAGATGGC | 293 |
| QY | 113 | AsnValIleAspCysValIlyAsnArgIleuValIlyGluGlyGluAspProGluNlyVal | 132 |
| Db | 294 | AAAGTGGCGAGCTGCTCAAGCGCGGCTCTGTCAGAGAGGCGACGACCCTCAAGGTG | 353 |
| QY | 133 | AlaAlaAspGluLeuValIlyAsnAspGlyValAspIleAsnIleAsnIleGlyIlyAspAsp | 153 |
| Db | 354 | GCGCGGACCAAGCTATCAAGACCGGCTGACGCTGCTCACAGATCGCGCGGACGAC | 413 |
| QY | 153 | ThrAsnThrAlaAlaIleAspLeuAlaIlePheLeuAlaArgAsnAsnTyrGlyLeuThr | 172 |
| Db | 414 | ACCAACACGGCGCGCGGACCTGGGGGGTACTCAAGGCGACGCGCTACACCTCGC | 473 |
| QY | 173 | ValIleGlyLeuProIleThrValIleAsnAspValPheProIleLysGluSerLeuGly | 192 |
| Db | 474 | GTAATCGGCTGCGCAAGACGATTCAGCAACGATTTGGCCCGTGGCGCACTGCTGGC | 533 |
| QY | 193 | AlaTyrThrAlaAlaGluGluGlyAlaArgTyrPheMetAsnValAlaGluAsnAsn | 212 |
| Db | 534 | GCCATGACGGCGCGGACGAGCGGCTGCGCTTCCAGAAAGTGTGCGGAGCAGACG | 593 |
| QY | 213 | AlaAsnProArgMetLeuIleValIleGluValMetGlyAsnGlyTyrPheLeuThr | 232 |
| Db | 594 | GCCAAACCGCGCTGCTGTGCTGCCACGAGGTCATGGGCGCCTCGTGGGTATCTGACG | 653 |
| QY | 233 | AlaAlaThrAlaGluGluTyrArgTyrIleLeuAspArgAlaGluTyrPheProGluIleu | 252 |
| Db | 654 | GCCCAAGCGCGGACATACACCGCGCGGACGCTCGCGCACCGGAGTTGCGGCCGACATC | 713 |
| QY | 253 | GlyLeuThrArgGluSerTyrGluValIleValIleValPheValProGluMetAlaIleAsp | 272 |
| Db | 714 | GCGCAACACCGCGAGCGCTACAGCATCCACGCGCTACGTGCGGAGATGACATCGAC | 773 |
| QY | 273 | LeuGluAlaGluAlaIlyAsnProLeuArgGluValIleMetAspLysValIlyAspCysValAsnIle | 292 |
| Db | 774 | CTCAAGCGGAGCGCGCGCTGCGCGCGCTCATAGAGCGCGGTGGCTGCTCAACATC | 833 |
| QY | 293 | PheValSerGluGlyAlaGlyValGluAlaIleValIleValIleGluMetGluAlaIleGlyGlu | 312 |
| Db | 834 | TTCTCTAGAGAGGCGCGGCGCATCAACGACATCGTCCGAGATGACCGCCACAGGCGAG | 893 |
| QY | 313 | GluValIleProArgAspAlaPheGlyHisIleLysIleLeuAspAlaValAsnProGlyLysTyr | 332 |

| Accession | Sequence | Position |
|-----------|---|----------|
| Dh | ACGCTCCCGCGCACCCCTTCGGCCACAGTCAGATCCATCAACCCCGGCGCTGG | 953 |
| Qy | PhGGIAGLInghpheaIagImeCIIeGIvAIGluYrthreuvAIGluYSeSeGIy | 352 |
| Dh | TTGGCGAAGAGATTCCGGCGCATGCTCGGCGGACAGAGTCTCTCTCCAGAGTCCGGC | 1013 |
| Qy | TyrPhehIAATgAIsaerAIsaerAanValAspAspMetrIgluIleIySeSeCyala | 372 |
| Dh | TACTTCTCGGCGTCGGCCCCCGCAACGCCGAGAGACTCCGCTCATCAAGGCGATGTG | 1073 |
| Qy | AspIeuAIAvaIGluCySaIaphaEArGArGIuSeCIyValIIeGIyHIsaSPGIuAaP | 392 |
| Dh | GACCTGGCCGTCCAGACTGCGCGCTGGCGGAGAGCTGTGCTGATCGGCCACGACGAGAGAG | 1133 |
| Qy | AaenGIyAanValIeuArGAIaIIeGIuPhPProArGIleIyS | 406 |
| Dh | CGCAACGGCTGCTCGCCCGCGATCGAGTTTGAGCCGCTCAAG | 1175 |

| RESULT 2 | CD379187 | LOCUS | DEFINITION |
|----------|--|-------|-----------------|
| CD379187 | 881 bp | mRNA | linear |
| PMW04335 | Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum tricornutum cDNA 5', mRNA sequence. | | EST 31-MAY-2003 |

| | |
|-----------|--|
| ACCESSION | CD379187 |
| VERSION | CD379187.1 |
| KEYWORDS | GI:31254801 |
| SOURCE | BST. |
| ORGANISM | Phaeodactylum tricornutum Phaeodactylum tricornutum <i>Phaeodactylum tricornutum</i> |

REFERENCE
AUTHORS
TITLE
JOURNAL
VOLUME
PAGE
YEAR

| | | | | | | | |
|---------|--|------------|-----|--------|-----|------|--------|
| MEDLINE | 22111123 | Hydrobiol. | 122 | (1977) | 222 | 1982 | (1982) |
| PUBMED | 12114555 | | | | | | |
| COMMENT | Contact: Bowler C Laboratory of Molecular Plant Biology Stazione Zoologica 'Anton Dohrn' | | | | | | |

Villa Comunale, I-80121, Napoli, Italy
Tel: 39 081 583 3268/3211
Fax: 39 081 764 1355
Email: christalpha@zn.it
Diatom EST Database (<http://aveschagen.szndwler.com>)

```

FEATURES
  source      Seq primer: T3 backward
              POLYA=yes.
              location/Qualifiers
              1..881
              /organism="Phaeodactylum tricornutum"

```

```

/mol_type="RNA"
/db_xref="taxon:2850"
/cell_line="C6MP632"
/clone_lib="Phaeodactylum tricorutum Uni-Zap XR"
/notes="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
Xho I"
```

| ORIGIN | ALIGN |
|-------------------|--------------|
| Alignment Scores: | |
| Pred. No.: | 1,41e-78 |
| Score: | Length: 881 |
| | Matches: 160 |
| | Matches: 160 |

| | | | |
|------------------------|--------|---------------|----|
| Percent Similarity: | 73.81% | Conservative: | 57 |
| Best Local Similarity: | 54.42% | Mismatches: | 65 |
| Query Match: | 32.22% | Indels: | 12 |
| DB: | 6 | Gaps: | 6 |

US-10-701-200-6 (1-437) x CD379187 (1-881)

QY 111 LeuThrAsnValLysAspCysValLysArgGlyLeuValysGluGluLysAspProGln 130
|||||
Db 3 CTACGAAAGCTTGCAGCTGCATTAAAGAGAGATCTCCAAACAGCGAGCAGCCCTTG 62
|||||

Oy 131 Ly6ValAlaAlaAspGlnLeuValLy6AspGlyValAspIleIleuHisThrIleGly 150

```

Db      63 GAAGTGGCTTCGAGACGATCATCAAGACCAAGTCCACCTCGTTCATCGATTGGCGGA
Qy      151 AAPPAGTTAAsnThrAlaAlaAspLeuAlaPheLeuAlaArgAsnAsnTyrGly
Db      123 GACGACACCAACACACAGAGTCTTACCTTTCCGACTCTGTGGAAAAGCAATGGG
Qy      171 ---LeuThrValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGln
Db      183 AAGTGGCGGTATCGTATGCCAAGACATCGACACATGTTATCTATTATGACAG
Qy      190 SerLeuGlyAlaTrpThrAlaAlaGluGlnGlyAlaArgTyrPheMetAsnValValAla
Db      243 ACCTTGGAGGCCACACAGCTGCGCGGGGGCGACGATTTTCAGAAAGTTGCAAT
Qy      210 GluAsnAspAlaAsnProArgMetLeuIleValHisGluValMetGlyArgAsnGly
Db      303 GAATGCACTGCCAATCTCGCATGCTTATTTCTCACAGATGATGGACCGGATGGGA
Qy      230 TrpLeuThrAlaAlaThrAlaGlnGluTyrArgLysLeuLeuAspArgAlaGluTrpLeu
Db      363 TACCTTACGGCGCTACGGCAGAAAGTACCGCATTTCTCAT---GCCCAAGACTTG
Qy      250 Pro-----GluLeuGlyLeuThrArgGluSer---TyrGluValHisAlaValPhe
Db      420 CCGCTGCGCTCGACCTTCCCTTCACCGCAAGTCTGCCGCGATTTATTCACGCAATTGG
Qy      266 ValProGluMetAlaIleAspLeuGluValGluValLysArgLeuArgLysValMetAsp
Db      480 ATTCGCCAATCTCAAGCTAGACTGTGTGCGAGTCCGCGCTTTGAAAGAAAGTCATAGAC
Qy      286 LysValAspCysValAsnIlePheValSerGluGlyValGlyValGluAlaIleValAla
Db      540 GAAGTGGTGGCTCAACATCTTTTCGGGAGAGGACCGGCTCGACGAAATCGTCGCG
Qy      306 GluMetGlnAlaLysGlyGlnGluValProArgAspAlaPheGlyHisIleLysLeuAsp
Db      600 GACATGAGAGCCCAACGCTGAGCGCGCGCGCGATTTGACACATCGATTTGGCT
Qy      326 AlaValAsnProGlyLysTrpPheGlyGluGlnPheAlaGluMetIleGlyAlaGluVal
Db      660 AANAATCAATCCCGACAGTACTTCTCCCAACACTTGCGGACAAATATGTGTCCGAAAG
Qy      346 ThrLeuValGlnLys---SerGlyTyrPhe---AlaArgAlaSerAlaSerAsn---ValAspAsp
Db      720 ACCATGTGCAAAAGTTCGGGATACCTTGGCCCGTTCGAGCGGCGCCACGATTTTCAT
Qy      365 MetArgLeuIle---LysSerCysAlaAsp---LeuAlaValGluCysAlaPheArgArg
Db      780 CGNCAACTCATCGGCGGCTGTCGCCAAGGAGTGGCGGTCGCGCGCTTTTGACNGGA
Qy      383 u---SerGlyValIleGlyHisAspGluAspAsn
Db      840 CATTCGAGTCTGTGGAAAGGATGAAAC

```

```

RESULT 3
LOCUS      CD383878      873 bp      mRNA      linear      EST 31-MAY-2003
DEFINITION PTFM09060 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
tricornutum cDNA 5', mRNA sequence.
ACCESSION  CD383878
VERSION     CD383878.1
KEYWORDS   GI:31259492
SOURCE      Phaeodactylum tricornutum
ORGANISM   Phaeodactylum tricornutum
Bukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 873)
Scale: S., Carels N., Falciatore A., Chliuano M.L. and Bowler C.
Genome properties of the diatom Phaeodactylum tricornutum
JOURNAL    Plant Physiol. 129 (3), 993-1002 (2002)
MEDLINE    2211123
PUBMED     12114555

```

```

COMMENT    Contact: Bowler C
            Laboratory of Molecular Plant Biology
            Stazione Zoologica 'Anton Dohrn'
            Villa Comunale I-80121, Napoli, Italy
            Tel: 39 081 583 3268/3211
            Fax: 39 081 764 1355
            Email: chris@alpha.szn.it
            Diatom EST Database (http://aves.echagen.szn.bowler.com)
            Seq primer: T3 backward
            POLY=A=Yes

FEATURES   Location/Qualifiers
            source          1..873
                        /organism="Phaeodactylum tricornutum"
                        /mol_type="mRNA"
                        /db_xref="taxon:2850"
                        /cell_line="CCMP632"
                        /clone_idb="Phaeodactylum tricornutum Uni-Zap XR"
                        /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
                        Xho I"

ORIGIN
Alignment Scores:
Pred. No.:      1,18e-75      Length:      873
Score:          712.00      Matches:      151
Percent Similarity: 73.96      Conservative: 45
Best Local Similarity: 56.98      Mismatches: 60
Query Match:    31.174      Indels:      9
DB:              6      Gaps:      4

US-10-701-200-6 (1-437) x CD383878 (1-873)
Qy      171 LeuThrValIleGlyLeuProLysThrValAspAsnAspValPheProIleLysGlnSer
Db      7 GTCGCGGTATCGGTATGCGCAAGACATCGACAGATGTTATCTATTATGACAGACC
Qy      191 LeuGlyAlaTrpThrAlaIleGluGlnGlyAlaArgTyrPheMetAsnValValAlaGlu
Db      67 TTCGAGCGGACACACAGCTCCGCGGAGGGGCGACGATTTTCAGAAAGCTGTCAATGAA
Qy      211 AsnAsnAlaAsnProArgMetLeuIleValHisGluValMetGlyArgAsnGlyTrp
Db      127 TGCATGCGCAATCTCTGATGCTTATCTTCACAGATGATGGAGCGCATTCGGGATAC
Qy      231 LeuThrAlaAlaThrAlaGlnGluTyrArgLysLeuLeuAspArgAlaGluTrpLeuPro
Db      187 CTTACGCGCGCTACGCGACAGAAAGTACCGGACATTTCTCAT---GCCAAGACTGCCC
Qy      251 -----GluLeuGlyLeuThrArgGluSer---TyrGluValHisAlaValPheVal
Db      244 GTCGCTCGGACTTGCCTTCCACAGAGTCTGCCGAGATTCACCGCATTTGGATT
Qy      267 ProGluMetAlaIleAspLeuGluAlaGluAlaLysArgLeuArgGluValMetAspLys
Db      304 CCGGAATCAAGCTGATCTGTGCGGAGTGGCGCGCTTTGAAAGAAAGTCATGACGAA
Qy      287 ValAspCysValAsnIlePheValSerGluGlyAlaGlyValGluValAlaIleValAlaGlu
Db      364 GTCGCTGCTCAACATCTTTTTCGGCGAGGCGACGGCGTGCAGAAATCTGCGGGAC
Qy      424 ATGAGAGCCACAGGTGAGACCGGCGCGCGCGATGCTTTGGACACCTCAGCTGCGTAAA
Qy      327 ValAsnProGlyLysTrpPheGlyGluGlnPheAlaGluMetIleGlyAlaGluValTrp
Db      484 ATCAATCCCGGACAGTACTTCTCCCAACCTTGCGGACAAATATGTGTCCGAAAAGACC
Qy      347 LeuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValAspAspMetArg
Db      544 ATGTGCAAAAGTCCGAGTACTTTCCTTCGCGCGCGCGCGCGCAAGATTTTCATGCCCA
Qy      367 LeuIleLysSerCysAlaAspLeuAlaValGluCysAlaPheArgArgLysSerGlyVal

```


Db 604 CTCATCGGGGCTGTGCGCGGCTGCGCCGCTATTGACGACATTCGGGATGC 663
 QY 387 TleeglyH1sAspG1uAspAenG1uAenVal---Leu-ArgAlaIleGluPheProArg11 405
 Db 664 ATGGGACGAGATGAGAAACAACCAACGACGATTTGAGCGATTTGATTCAGTCGGAT 723
 QY 405 e-LyeglyG1uYlyPheProPheAen11eAspThrAspTTPheAenSerMetLeuSer-Glu 424
 Db 724 CAAGAGGNGGCAACCTTTGATTTCTCAGGAATGGGTCCNACAAATGCTCAAGGGA 783
 QY 425 TleeglyGln 427
 Db 784 ATTGACAA 792

RESULT 4
 BM321461
 LOCUS 743 bp mRNA linear EST 03-JAN-2002
 DEFINITION Mestigamoeba balamuthi lambda ZAP II library
 Mestigamoeba balamuthi cDNA similar to pyrophosphate-fructose
 6-phosphatase 1-phosphotransferase (EC 2.7.1.9), mRNA sequence.

ACCESSION BM321461 GI:18055867
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mestigamoeba balamuthi
 Eukaryota; Pelobiontida; Mestigamoebidae; Mestigamoeba.
 1 (bases 1 to 743)
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
 Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
 Philippe,H.
 The analysis of 100 genes supports the grouping of three highly
 divergent amoebae: Dictyostelium, Entamoeba, and Mestigamoeba
 Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Muller M&Kios
 Laboratory of Biochemical Parasitology
 The Rockefeller University
 1230 York Avenue, New York, NY 10021, USA
 Email: mmuller@rockefeller.edu
 Insert Length: 743 Std Error: 0.00
 POLYA=Yes.

FEATURES
 Source
 1..743
 /location="Qualifiers
 /organism="Mestigamoeba balamuthi"
 /mol_type="mRNA"
 /strain="ATCC 30984"
 /db_xref="taxon:108607"
 /clone_lib="Mestigamoeba balamuthi lambda ZAP II library"
 /note="syn: Phreatamoeba balamuthi"

ORIGIN
 Alignment Scores:
 Pred. No.: 9,42e-71 Length: 743
 Score: 671.00 Matches: 126
 Percent Similarity: 80.00% Conservative: 30
 Best Local Similarity: 64.62% Mismatches: 39
 Query Match: 29.38% Indels: 1
 DB: 4 Gaps: 0

US-10-701-200-6 (1-437) x BM321461 (1-743)

QY 243 LeuAspArgAlaGluTrpLeuProGluLeuGlyLeuThrArgGluSerTyrGluValHis 262
 Db 1 CTCGGGCAACCGGAGATTGGCGCCGAGCTCGGCGACAGCGCGGAGCGCTACGACATCCAC 60
 QY 263 AlaValPheValProGluMetAlaIleAspLeuGluAlaGluAlaValSerLeuArgGlu 282
 Db 61 GCGGCTACGTCGCCGAGATGACATCGACCTCAAGCGGAGGCGCCCGCTGCGCGCC 120
 QY 283 ValMetAspLeuValAspCysValAen11ePheValSerGluG1uAlaGluValGluAla 302
 Db 121 GTCAATGAGCGGCTGGCTGCGTCAACATCTTCTCAGCGAGGCGCGGCGCATCAACGAC 180

QY 303 TleValAlaGluMetGlnAlaValGlyGlnGluValProArgAspAlaPheGlyHisIle 322
 Db 181 ATCTCGCGGAGATGACCGCCAGGCGAGCGGCTCCGCGGACCC-TTCGGCACAGTC 239
 QY 323 LysLeuAspAlaValAenProGluYlySerTrpPheGlyGluGlnPheAlaGluMetIleGly 342
 Db 240 AAGATCGACCTTCATCAACCCCGCGCTGTTCGGCAACAGTTTCGGCGGCGATGGTCGGC 299
 QY 343 AlaGluYlyThrLeuValaGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAenVal 362
 Db 300 GCGGACAAAGTCCTGCTCAGAAAGTCGGGCTACTTCTCGCCCTGCGCCCGCCAAACGCC 359
 QY 363 AspAspMetArgLeuIleLysSerCysAlaAspLeuAlaValaGluCysAlaPheArgArg 382
 Db 360 GAGGACCTCCCGCTTCATCAAGCGCATGGGTGACCTGGCGGCTGACCTGCGCGCGC 419
 QY 383 GluSerGlyValIleGlyH1sAspG1uAspAenG1uAenValIleuArgAlaIleGluPhe 402
 Db 420 GAGGCTGCGCTGATCGGCGCACGAGGAGCGCAACGCGCGTCTGCGCGCATCGATTT 479
 QY 403 ProArgIleLysGlyGlyLysPheProPheAen11eAspThrAspTTPheAenSerMetLeu 422
 Db 480 GAGCGCGTCAAGGCGGCGCAAGCGCTTCAACATCGACCAACCGGTGTTCAACCACTCTTC 539
 QY 423 SerGluIleGlyGlnProLysGlyGlyLysValaGluValSerHis 437
 Db 540 AACGAGATCGGCGACGCGCCAAAGGCGCGCAAGGTCTCGTCCGCCAC 584

RESULT 5
 CD379057
 LOCUS 896 bp mRNA linear EST 31-MAY-2003
 DEFINITION PTM04226 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 tricornutum cDNA 5', mRNA sequence.

ACCESSION CD379057
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 896)
 Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 Genome properties of the diatom Phaeodactylum tricornutum
 Plant Physiol. 129 (3), 993-1002 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.zn.dohrn.it
 Diatom EST Database (http://aves.thagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=Yes.

FEATURES
 Source
 1..896
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Xho I"

ORIGIN
 Alignment Scores:
 Pred. No.: 5,89e-70 Length: 896
 Score: 665.50 Matches: 142
 Percent Similarity: 73.83% Conservative: 47
 Best Local Similarity: 55.47% Mismatches: 57

QY 391 GUAAPANGLYASNAVal---LeuARGAlaIleGluPheProArgIleValGlyGlyLeu 409
 DB 599 GAAGGACAAACCCAAACAGCCCATTCGAGGATGTAATTCATGCAATCAAGGGTGCCAA 658
 QY 410 -ProheanleleapThraP-TpPheanSerMetLeuSer-GluIleGlyGln 427
 DB 659 CCCCTTTGATTTTCTCANGAAATGGNTTCCCAAAATGCTCAAGGAAATTTGGACAA 715

RESULT 7
 CD383585 878 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM008757 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD383585
 VERSION CD383585.1 GI:31259199
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 878)
 AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diction EST Database (http://aveschagen.szbwler.com)
 Seg primer: T3 backward
 POLYA=yes

FEATURES
 source Location/Qualifiers
 1..878
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN
 Alignment Scores: 1.15e-61 Length: 878
 Pred. No.: 597.50 Matches: 129
 Score: 70.894 Conservative: 39
 Percent Similarity: 54.434 Mismatches: 61
 Best Local Similarity: 26.164 Indels: 8
 Query Match: 6 Gaps: 5
 DB: 6

US-10-701-200-6 (1-437) x CD383585 (1-878)

QY 70 TYRARGLYLYTYRILEuLeuLeuLeuGlyAspSerTy-ProValThrAlaGluVal 89
 DB 62 TACCTTGGGGGTACAAAGAAATGCTTACGGAGATTCATGCTATTCGGAATCA 121
 QY 90 ARGLYSLYALIGLYValLeuGlnArgPheGlyGlySerValIleGlyAsnSerArgVal 109
 DB 122 TGGGATTCACTCGACTCTCTGAATACGGTGGAGATCGCCCATTTGGAATTCGCGCGTC 181
 QY 110 LYSLEuThrAsnValLYSAspCySValLYSARGLYLeuValLYSGLuGlyGluAspPro 129
 DB 183 AAGCTGACCAACGGTGCATTCGATTAAGAAGGGATACGTCACACGCGACGACGCC 241
 QY 130 GlnLYSValAlaAlaAspGlnLeuValLYSAspGlyValAspIleLeuHisThrIleGly 149
 DB 242 TTAGAAGTACCTTACACAGACTCATCAAGACCAAGTCCAGCTGCTTACATCGATTGGC 301

QY 150 GLYAspAspThrAsnThrAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTy 169
 DB 302 GAGACGACACCAACACACAAAGCTGACCTTTCCGACTTCGAGAAAGACAAAT 361
 QY 170 GLY---LeuThrValIleGlyLeuProLYSThrValAspAsnAspValPheProIleLYS 188
 DB 362 GGCAAGTCGCGCTCATCGATATGCCCCAAGACATCGACAAAGATGTTATCTATTCG 421
 QY 189 GlnSerLeuGlyAlaATPThrAlaAlaGluGlnGlyAlaArgTYRPhMetAsnValVal 208
 DB 422 CAGACCTTCGAGGCGACACAGCTGCCCTACAGGGGCGACCGCTTTTCGAGAACGTTGTC 481
 QY 209 AlaGluAsnAlaAsnProArgMetLeuIleValHisGluValMetGlyArgAsnCyS 228
 DB 482 AATGAACTGACGCCAAATCTCGCATGCTTATTCCTCCAGAGTGATAGAGCCATTCG 541
 QY 229 GLYTrpLeuThrAlaAlaThrAlaGlnGluTYRArgLYSLeuLeuAspArgAlaGluTrp 248
 DB 542 GGATTAACCTTACGGCGGCTACCGGACAAAGTACCGGCACATTCTCAAC---GCCAAGC 598
 QY 249 LeuPro-----GluLeuGlyLeuThrArgGlySer---TYRGLuValHisAlaVal 264
 DB 599 TTGCCCGTCGCGCTCGACTTCGCCCTTCACCGCAAGTCTGCCCGGATATTCACGCCAT 658
 QY 265 -PheValProGluMetAlaIleAspLeuGluAlaGluAlaLYSArgLeuArgGluValMe 284
 DB 659 TTGGATTCCCAACTCAAGCTTAACTTGCTGCGCGCAAGTCTGCGCTTTGANAANAAGTCAT 718
 QY 284 CASpLYSValAspCySValAsnIle---PheValSerGluGlyAlaGly 299
 DB 719 GGGCGAACTCGGTTCCGTTAACATCTTTTTCNGGAAAGGACCGGCG 767

RESULT 8
 CD382389 904 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM007559 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD382389
 VERSION CD382389.1 GI:31258003
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 904)
 AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diction EST Database (http://aveschagen.szbwler.com)
 Seg primer: T3 backward
 POLYA=yes

FEATURES
 source Location/Qualifiers
 1..904
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN
 Alignment Scores: 1.74e-60 Length: 904
 Pred. No.: 1

Score: 588.00 Matches: 131
 Percent Similarity: 71.84% Conservative: 45
 Best Local Similarity: 53.47% Mismatches: 59
 Query Match: 25.74% Indels: 10
 DB: 6 Gaps: 4

US-10-701-200-6 (1-437) x CD382389 (1-904)

QY 157 AAlaAlaAspLeuAlaAlaPheLeuAlaArgAsnAsnTyrGly---LeuThrValIleGly 175
 DB 2 GGTGTAACCTTCCGACTACCTGTGGAAAGACAAATGGCAAAATCGCGCTCATCGGT 61
 QY 176 LeuProLyThrValAlaPheAsnAspValPheProIleLysGlnSerLeuGlyAlaTPrThr 195
 DB 62 ATGCCCAAGACCACTCACAACGATGTTATCCATTTGTCAGACCTTCGAGCGGACACA 121
 QY 196 AAlaAlaGlnGlnGlyAlaArgTyrPheMetAspValAlaGlnLysAsnAlaAsnPro 215
 DB 122 GCTGCCGTACAGGGGGGACGCTTTTCGAGAACGTTGCAATGATGACACTGCCAATCCT 181
 QY 216 ArgMetLeuIleValIleGlnValMetGlyArgAsnCyGlyTyrPheThrAlaAlaThr 235
 DB 182 CGGATCCTTATTTCCACAGGTGTATGGACGCGATTCGGGATACCTTACGCGCTTAG 241
 QY 236 AAlaGlnGlyTyrArgLysLeuLeuAspArgAlaGlnTPrLeuPro-----GluLeu 252
 DB 242 GCACAAAGTACCGCACATTTCTCAAC---GCCCAAGCTTCGCTCGCTCGACTTG 298
 QY 253 GlyLeuThrArgGlnSer---TyrGlnValHisAlaValAlaPheValProGlnMetAlaIle 271
 DB 299 CCTTCCACCGGACAGTCTGCCCGCGATTCACGCAATTTGATCCGAACTCAAGCTA 358
 QY 272 AspLeuGlnAlaGlnValLysArgLeuArgGlnValMetAspLysValAspCyValAsn 291
 DB 359 GACTGTGTGGCGGAGTGGCGGCGCTTGAGAAAGTATGACAGCAATCCGCTGCGTCAAC 418
 QY 292 IlePheValSerGlnGlyAlaGlyValAlaIleValAlaGlnLysGlnAlaLysGly 311
 DB 419 ATCTTTTTCGGCGGACGCGGCGCTGCANGAATGCTGCGGACATGAGGCGCAATGT 478
 QY 312 GlnGlnValProArgAspAlaPheGlyHisIleLysLeuAspAlaValAsnProGlyLys 331
 DB 479 GACGCGGTCCGCGGATGCTTTGACACGTCACCTGCTAAATCAATCCCGAGACAG 538
 QY 332 TrpPheGlnGlyGlnPheAlaGlnMetIleGlyAlaGlnLysThrLeuValGlnLysSer 351
 DB 539 TTCTTCTCCCAACACTGCGCGGACATATTTGTCGCGGAAAGCCATTTGTGCAAAAGTCG 598
 QY 352 GlyTyrPheAlaArg-AlaSerAlaSerAsn-ValAspAspMetArgLeuIle-LysSer 370
 DB 599 GGATACCTTTGGCGGTTCCGACGCGGCAACNAATTTTGATCGCCCAATCATCGCGGCGC 658
 QY 371 CyeAla-AspLeuAlaValGlnCysValAlaPheArgGlnSerGlyValIleGlyHisAs 390
 DB 659 TGTNCCGANGCTGGCGTNGCGCGCTTATTTGACNAGAACTTCCGNAATTCNTGGGA 718
 QY 390 pGluAsp 392
 DB 719 CAAGAT 725

RESULT 9
 LOCUS CD378198 932 bp mRNA linear EST 31-May-2003
 DEFINITION PTM00367 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD378198
 VERSION CD378198.1 GI:31253812
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 REFERENCE 1 (bases 1 to 932)

AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 TITLE Genome properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 22111123
 PUBMED 12114555

COMMENT
 CONTACT: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Dicom EST Database (http://aveslabgen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source Location/Qualifiers

1..932
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Kpn I"

ALIGNMENT SCORES:

Pred. No.: 1,02e-57 Length: 932
 Score: 555.50 Matches: 115
 Percent Similarity: 77.27% Conservative: 38
 Best Local Similarity: 56.08% Mismatches: 39
 Query Match: 24.76% Indels: 6
 DB: 3 Gaps: 3

US-10-701-200-6 (1-437) x CD378198 (1-932)

QY 164 LeuAlaArgAsnAsnTyrGlyLeuThrValIleGlyLeuProLyThrValAspAsnAsp 183
 DB 31 TTGGAAGAACACAAATGGCAAAATCGCGATCGGATGCCAAGACCATCGCAACGAT 90
 QY 184 ValPheProIleLysGlnSerLeuGlyAlaTPrThrAlaAlaGlnGlnAlaArgTyr 203
 DB 91 GTTATTCCTATTGTACAGACCTTCGAGCGGACACAGCTGCGTGCAGGGGACACATTT 150
 QY 204 PheMetAspValAlaGlnLysAsnAlaAsnProArgMetLeuIleValHisGlnVal 223
 DB 151 TTGGAAGACGTTGATCAATGACACTGCAATCCTCGCATCTTATTTCCACGAGGT 210
 QY 224 MetGlyArgAsnCyGlyTyrPheThrAlaAlaThrAlaGlnGlnLysArgLysLeu 243
 DB 211 ATGGGACGGGATTCGGGATACCTTACGGCGCTAGCGGACAAAGTACCGGACATTTCTC 270
 QY 244 AspArgAlaGlnTPrLeuPro-----GluLeuGlyLeuThrArgGlnSer---Tyr 259
 DB 271 AAT---GCCCAAGACTTGGCGCTGCGCTGCGCTGCGCTTCCACGCAAGTGTGCCGC 327
 QY 260 GlnValHisAlaValAlaPheValProGlnMetAlaIleAspLeuGlnAlaGlnLysArg 279
 DB 328 GATATTCAGCCATTGATTCCTCCGAACTCAAGCTAGATGATGTGCGGAGTGGCGGT 387
 QY 280 LeuArgGlnValMetAspLysValAspCyValAsnIlePheValSerGlnGlyAlaGly 299
 DB 388 TTGAAGAAAGTCATGAGACGAACTCGGTTGCGTCAACATCTTTTTCGCAAGGACCGCG 447
 QY 300 -ValGlnAlaIleValAlaGlnMetGlnAlaLysGlnGlnValProArgAspAlaPh 319
 DB 448 CGTGCAGAAATCGTCGCGGACATGAGGCGCAACGNGAACCGTGCCTGCGGAGTCCCT 507
 QY 319 eGlyHisIleLysLeuAspAlaValAsnProGlyLysTPrPheGlyGlnGlnPheAlaG 339
 DB 508 TGGACGCTCACGTTGGCTAAATCAATCCGGAAGTACTTCTCCCAACACTTGGCGGA 567
 QY 339 mMetIleGlyAlaGlnLysThrLeuValGlnLysSerGlyTyrPheAlaArg 356

```

Db      568 CAATATTGCTGCCGAAAGAACCATTTGTGCAAAAGTCGGATCTTGGCCGT 619
RESULT 10
LOCUS   CD384626                      746 bp      mRNA      linear      EST 31-MAY-2003
DEFINITION tricornutum cDNA 5', mRNA sequence.
ACCESSION CD384626
VERSION   CD384626.1 GI:31260240
KEYWORDS EST.
SOURCE    Phaeodactylum tricornutum
ORGANISM Phaeodactylum tricornutum
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
REFERENCE 1 (bases 1 to 746)
AUTHORS   Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
TITLE      Genome properties of the diatom Phaeodactylum tricornutum
JOURNAL    Plant Physiol. 129 (3), 993-1002 (2002)
MEDLINE    2211123
PUBMED     12114555
COMMENT    Contact: Bowler C
            Laboratory of Molecular Plant Biology
            Stazione Zoologica 'Anton Dohrn'
            Villa Comunale, I-80121, Napoli, Italy
            Tel: 39 081 583 3268/3211
            Fax: 39 081 764 1355
            Email: chris@alpha.szn.it
            Diatom EST database (http://aves.thagen.sznbowlr.com)
            Seq primer: T3 backward
            POLYA=No.

FEATURES
     source             1..746
                        /organism="Phaeodactylum tricornutum"
                        /mol_type="mRNA"
                        /db_xref="taxon:2850"
                        /cell_line="CCMP632"
                        /clone_11b="Phaeodactylum tricornutum Uni-Zap XR"
                        /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
                        Xho I"

ORIGIN
Alignment Scores:
Pred. No.:          1,11e-57      Length:          746
Score:              564.00        Matches:          113
Percent Similarity: 73.53%        Conservative:      37
Best Local Similarity: 55.39%      Mismatches:        48
Query Match:        24.69%        Indels:           6
DB:                  6            Gaps:              4

US-10-701-200-6 (1-437) x CD384626 (1-746)

QY      229 GATTTPLLeuThAlaIaThraIaGInGluTyrArGlyLeuLeuAaPArGAlaGluTrp 248
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9 GGATACCTTAACGGCCGCTACGGCAAAAGTACCGGACATCTCAAC---GCCCAAGGC 65
QY      249 LeuPro-----GluLeuGlyLeuThraTArGluSer--TyrGluValHisAlaVal 264
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      66 TTGCCCGTGGCTCGGACTTGGCCCTTCCACCGCAAGTCTGCCCGCGATATTCACGCGCAT 125
QY      265 PheValProGluMeValaIleAePLeuGluValaGluValaLbLbAePLeuArGluValaMet 284
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      126 TGGATTCCCGAACTCAAGCTAGACTTGTCGCGCGAGTCCGCGCTTGAAGAAAGTCATG 185
QY      285 AaPLeValaAaPLeuValaAaPLeuValaAaPLeuValaAaPLeuValaAaPLeuValaAaPLeu 304
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      186 GACGAAGTGGCTTGGCTCAACATCTTTTCGGCGAGGAGGACCGGCGCTGCAGCAATCTGTC 245
QY      305 AAlaGluMeGAlaAlaLysGlyGlnGluValaProArGAsPAlaPheGlyHisIleIleuLeu 324
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      246 GCGGACATGAGAGGCCAATGATGTGACGCGCGCGGATGCTTGGACACATCAGCTTG 305
QY      325 AaPAlaValaAaPLeuProGlyLeuTrrPheGlyGluGlnPheAlaGluMetIleGlyAlaGlu 344

```

```

Db      306 GCTAAATCAATCCCGGACACTACTTCTCCCAACACTTGGGCGACAAATATTGTCGCCA 365
QY      345 LysThrIleuValGlnLysSerGlyTyrPheAlaArgAlaSerAlaSerAsnValaAaPLeu 364
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      366 AAGACCATTTGCAAAAGTCGAGTACTTTCGCCGTTCCGACGCGGCAACGATTTGCAT 425
QY      365 MetArgLeuIleLysSerCysAlaAaPLeuAlaValGluCysAlaPheArgArgGluSer 384
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      426 CGCCAACTCATCGGGGCTGTGCGGAGCTGCGCTGCGCGCTGATTTGACGAGCATTTCC 485
QY      385 GlyValIleGlyHisAePLeuAaPLeuValaAaPLeuValaAaPLeuValaAaPLeuValaAaPLeu 403
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      486 GGATTCATCGGACAGATGAAGCAAAACCAACGCGCCATTCAGCGCATTTGATTCAGT 545
QY      404 ArgIleuLysGlyLysProPheAsnIleAaPLeuThraPLeuThraPLeuThraPLeuThraPLeu 423
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      546 CGCATCAAGGATGGCAACCTTGATATTCTCANGAATGATGTTCAACAATGCTCAAN 605
QY      424 GluIleGlyGln 427
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      606 GAATTTGACAA 617

RESULT 11
LOCUS   CD375650                      757 bp      mRNA      linear      EST 31-MAY-2003
DEFINITION tricornutum cDNA 5', mRNA sequence.
ACCESSION CD375650
VERSION   CD375650.1 GI:31251264
KEYWORDS EST.
SOURCE    Phaeodactylum tricornutum
ORGANISM Phaeodactylum tricornutum
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
REFERENCE 1 (bases 1 to 757)
AUTHORS   Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
TITLE      Genome properties of the diatom Phaeodactylum tricornutum
JOURNAL    Plant Physiol. 129 (3), 993-1002 (2002)
MEDLINE    2211123
PUBMED     12114555
COMMENT    Contact: Bowler C
            Laboratory of Molecular Plant Biology
            Stazione Zoologica 'Anton Dohrn'
            Villa Comunale, I-80121, Napoli, Italy
            Tel: 39 081 583 3268/3211
            Fax: 39 081 764 1355
            Email: chris@alpha.szn.it
            Diatom EST database (http://aves.thagen.sznbowlr.com)
            Seq primer: T3 backward
            POLYA=No.

FEATURES
     source             1..757
                        /organism="Phaeodactylum tricornutum"
                        /mol_type="mRNA"
                        /db_xref="taxon:2850"
                        /cell_line="CCMP632"
                        /clone_11b="Phaeodactylum tricornutum Uni-Zap XR"
                        /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
                        Xho I"

ORIGIN
Alignment Scores:
Pred. No.:          3.62e-55      Length:          757
Score:              543.50        Matches:          122
Percent Similarity: 73.54%        Conservative:      42
Best Local Similarity: 54.71%      Mismatches:        48
Query Match:        23.80%        Indels:           11
DB:                  6            Gaps:              5

US-10-701-200-6 (1-437) x CD375650 (1-757)

QY      149 GtGtGtAaPArPThraThraIaIaIaAaPLeuAlaIaPhe---LeuAlaArgAaen 167

```

```

Db      1 GGGGAGAGACACACACACACAGCTGCTACCTTCGACACTGTTGAGAAAAGCAC 60
Qy      168 Aen-TyrglyLeuThrValIleGlyLeuProlysthrValAspAsnAspValPheProI 187
Db      61 NATGGCAAGAGATGCGCGTATCGGTATGCCCAAGACATCGACACGATTTATCTTAT 120
Qy      187 elyGlnserLeuGlyValATrThrAlaIleGlnGlyValArgTy-PheMetAsnVa 207
Db      121 TGTACAGACTTCGGAGCGGACACAGCTGCGGAGGAGGACGATTTTCAGAAACGT 180
Qy      207 ValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHleGluValMetGlyArgAs 227
Db      181 TGTCAATGATGACATGCGCAATCTCGCATGCTTATTCACAGATGTATGAGACGCA 240
Qy      227 ncyGlyTrpLeuThrAlaIleThrAlaIleGlnGlyValArgTy-PheMetAsnVa 247
Db      241 TTTGGGATACCTTACCGCGCTACCGGACAGAAATGACCGGACATTTCTCAAT--GCCA 297
Qy      247 uTrpLeuPro-----GluLeuGlyLeuThrArgGlySer--TyrglyValHleAl 263
Db      298 AGACTTGGCGCGTGGCTCGGACTTGCCCTTCCACCGCAAGTGTGCCCGGATTTACCG 357
Qy      263 aValPheValProGluMetAlaIleAspLeuGluAlaGluAlaValArgLeuArgGluVa 283
Db      358 CATTTGGATTCCTCCGAACTCAAGCTAGACTGTGCGCGAGTCCGCGCGCTTTGAAAGAA 417
Qy      283 lMetAspLysValAspCysValAsnIlePheValSerGluValAlaGluValAlaIle 303
Db      418 CATGACGAAAGTGTGGTGGTTCACATCTTTTCCGCGGAGGACGCGGCTGCAAGAAAT 477
Qy      303 eValAlaGluMetGlnAlaIleGlyGlnGluValProArgAspAlaPheGlyHleIleTy 323
Db      478 CGTCCGGGACATGAGAACCAAGCTGAACCGTCCGCGGAGTCCCTTGGACACGTCAAC 537
Qy      323 elMetAspAlaValAsn-ProGlyLysTrpPheGlyGluGlnPhe---AlaGluMetIleG 342
Db      538 GTTGGCTTAAATCAATNCCGGGACAGTACTTCTTCAACACTTNGCGGACATATTTGG 597
Qy      342 lValGluIleuThrLeu-ValGlnLysSerGlyTy-PheAlaArgAlaSerAlaSerA 361
Db      598 NGCCCAAAACCATTTGNGCAGAAAGTGGGATCTTTGCGCGTCCCGGCGGACAA 657
Qy      361 n 361
Db      658 C 658

```

RESULT 12
CD380358 639 bp mRNA linear EST 31-MAY-2003
LOCUS PTM005527 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
DEFINITION tricornutum cDNA 5', mRNA sequence.
ACCESSION CD380358
VERSION CD380358.1 GI:31255972
KEYWORDS EST.
SOURCE Phaeodactylum tricornutum
ORGANISM Phaeodactylum tricornutum
Bukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 639)
SCALA,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
Genome properties of the diatom Phaeodactylum tricornutum
Plant Physiol. 129 (3), 993-1002 (2002)
2211123
12114555
Contact: Bowler C
Laboratory of Molecular Plant Biology
Stazione Zoologica 'Anton Dohrn'
Villa Comunale, I-80121, Napoli, Italy
Tel: 39 081 583 3268/3211
Fax: 39 081 764 1355
Email: christel@alpha.szn.it
Diatom EST Database (http://aveethagen.sznbowl.com)

```

FEATURES
source
Seq primer: T3 backward
POLYA=No.
Location/Qualifiers
1..639
/organism="Phaeodactylum tricornutum"
/mol_type="mRNA"
/db_xref="taxon:2850"
/cell_line="CMP632"
/clone_1ib="Phaeodactylum tricornutum Uni-Zap XR"
/notes="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
Xho I"

```

ORIGIN

Alignment Scores:

| Pred. No.: | 5.54e-52 | Length: | 639 |
|------------------------|----------|---------------|-----|
| Score: | 516.50 | Matches: | 96 |
| Percent Similarity: | 74.57% | Conservative: | 33 |
| Best Local Similarity: | 55.49% | Mismatches: | 43 |
| Query Match: | 22.61% | Indels: | 1 |
| DB: | 6 | Gaps: | 1 |

US-10-701-200-6 (1-437) x CD380358 (1-639)

```

Qy      256 ArgGlySerTyrglyValHleAlaValPheValProGluMetAlaIleAspLeuGluAla 275
Db      23 AAGTGTGCGCGGAGATTCACGCCATTTGATTCGGAATCAAGCTAGACTGATGTCGCC 82
Qy      276 GluAlaLysArgLeuArgGluValMetAspLysValAspCysValAsnIlePheValSer 295
Db      83 GAGTCCGCGCGCTTGAAGAAAGTCAAGAGTGGTGGTTCACATCTTTTTCGCC 142
Qy      296 GluGlyAlaGlyValAlaGluAlaIleValAlaGluMetGlnAlaIleGlyGlnGluValPro 315
Db      143 GAGGGGACGGGGGTGAGAAATCGTCCGACATGTGAGGCCAAGCTGAAGCGCGCGG 202
Qy      316 ArgAspAlaPheGlyHleIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGlu 335
Db      203 CCGGATGCTTTGGACACGTCAAGTGGCTTAAATCAATCCCGGACAGTACTTCTCCAA 262
Qy      336 GluPheAlaGluMetIleGlyValGluIleuThrLeuValGlnLysSerGlyTy-PheAla 355
Db      263 CACTTGGCGGACAAATTTGTGCTCCGAAAGACATTTGCAAAAGTGGGATTCCTTGGC 322
Qy      356 ArgAlaSerAlaSerAsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAla 375
Db      323 CGTCCGCGCGCGGCGCAAGATTTTCATGCCCAATCATGGGGCGCTGTCCGAGCGTGGC 382
Qy      376 ValGluCysAlaPheArgArgGlySerGlyValIleGlyHleAspGluAsnGlyAsn 395
Db      383 GTCCGCGCGCGCTATTGACGGAATTCGCGATGTCAGAGACAGATGACAAACCAAC 442
Qy      396 Val---LeuArgAlaIleGluPheProArgGlyLeuGlyGlyLysTrpPheAsnIleAsp 414
Db      443 ACGCCCATTCGAGCAATTAATTCAGTCCGACCAAGGGGAGGAAACCTTTGATATTCT 502
Qy      415 ThrAspTrpPheAsnSerMetLeuSerGluIleGlyGln 427
Db      503 CANGAATGTTCCACAAATGCTCAAGAAATTTGACAA 541

```

RESULT 13
CD379504 916 bp mRNA linear EST 31-MAY-2003
LOCUS PTM004673 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
DEFINITION tricornutum cDNA 5', mRNA sequence.
ACCESSION CD379504
VERSION CD379504.1 GI:31255118
KEYWORDS EST.
SOURCE Phaeodactylum tricornutum
ORGANISM Phaeodactylum tricornutum
Bukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 916)
SCALA,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.

TITLE Genome properties of the diatom *Phaeodactylum tricornutum*
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: christ@alpha.szn.it
 Diatom EST Database (<http://aves.thgen.sznbowler.com>)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source 1..916
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_id="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN

Alignment Scores:

| Pred. No.: | 2,23e-51 | Length: | 916 |
|------------------------|----------|---------------|-----|
| Score: | 513.50 | Matches: | 96 |
| Percent Similarity: | 73.99% | Conservative: | 32 |
| Best Local Similarity: | 55.49% | Mismatches: | 44 |
| Query Match: | 22.48% | Indels: | 1 |
| DB: | 6 | Gaps: | 1 |

US-10-701-200-6 (1-437) x CD379504 (1-916)

```

OY 256 ArgGluSerTyrGluValHisAlaPheValProGluMetAlaIleAspLeuGluAla 275
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 11 AAGTCGCGCGGAGATATTCAGCCATTGTGATCCGGAATCAAGTAACTGATGTCGCGC 70
OY 276 GUAAlaValArgLeuArgGluValMetAspLysValAspCysValAsnIlePheValSer 295
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 71 GAGTCGCGCGCGCTTGAAGAAGATCATGACGAAGTCGCTGCTCAACATCTTTTTCGCGC 130
OY 296 GluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValPro 315
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 131 GAGGCGACGCGCGGTGCGAAGAAATCGTCGCGGACATGAGGCGCAATGTGTACCGCGCGC 190
OY 316 ArgAspAlaPheGlyHisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGlu 335
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 191 CGCGATGCTTTGGACACGTCACTGGCTAAATCAATCCCGGACAGTACTTCTCCCA 250
OY 336 GlnPheAlaGluMetIleGlyAlaGluLysTrpLeuValGlnLysSerGlyTrpPheAla 355
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 251 CACTTGCGCGGACAAATTTGTGCTCCGAAAGACCATTTGTCCAAAGTCGGGATATCTTTCGC 310
OY 356 ArgAlaSerAlaSerAsnValAspPheMetArgLeuIleLysSerCysAlaIleAspLeuAla 375
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 311 CGTTCGCGGACGCGCAACGATTTCCATCGCAATCATCGGGCGCTGTCCGAGGCTGGC 370
OY 376 ValGluCysAlaPheArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsn 395
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 371 GTGCGCGCGCGCTATTGACGAGCATTTCCGATCATGGGACAGAGATGAAACAAACCCCAAC 430
OY 396 Val----LeuArgAlaIleGluPheProArgIleLysGlyGlyLysTrpPheAsnIleAsp 414
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 431 AGCCCATTCGAGCGCATTTGAATTCAGTCCGATCAAGGGTGGCAACCTTTGATATTCT 490
OY 415 ThrAspTrpPheAsnSerMetLeuSerGluIleGlyGln 427
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 491 CAGGAATGTTCCAAACAATGCTCAAGAAATTGACAA 529
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

RESULT 14
 CD378314

LOCUS CD378314 566 bp mRNA linear EST 31-MAY-2003
 DEFINITION PTM003483 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD378314
 VERSION CD378314.1 GI:31253928
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; Eukaryotes; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 566)
 AUTHORS Scala,S., Carels,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
 TITLE Genome properties of the diatom *Phaeodactylum tricornutum*
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: christ@alpha.szn.it
 Diatom EST Database (<http://aves.thgen.sznbowler.com>)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source 1..566
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_id="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"

ORIGIN
 Alignment Scores:
 Pred. No.: 8,82e-51 Length: 566
 Score: 506.00 Matches: 103
 Percent Similarity: 74.47% Conservative: 37
 Best Local Similarity: 54.79% Mismatches: 42
 Query Match: 22.15% Indels: 6
 DB: 6 Gaps: 4

US-10-701-200-6 (1-437) x CD378314 (1-566)

```

OY 166 ArgAsnAsnTrpGlyLeuThrValIleGlyLeuProLysThrValAspAsnAspValPhe 185
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 5 AAGCAACAATGGCAAGTGCCTGATCGGTATGCCCAAGACCATCGACAAACGATGTTAT 64
OY 186 ProIleLysGlnSerLeuGlyValATrPThrAlaIleGluGlnGlyValArgTyrPheMet 205
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 65 CCTATTGTCAGACCTTGGAGCGGACACACACTGCGCGATAGAGGGGACGCTTTTCCAG 124
OY 206 AsnValValAlaGluAsnAsnAlaAsnProArgMetLeuIleValHisGluValMetGly 225
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 125 AACGTTGCATGAAATGACGACCAATCTCGCATGCTTATTCTCCAGAGTATGGGA 184
OY 226 ArgAsnCysGlyTrpLeuThrAlaIleThrAlaGlnGlyTyrArgGlyLeuLeuAspArg 245
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 185 CGCGATTGGGATACCTTACGCGCGCTACGACAAAGTACCGGACATCTTCAAC 241
OY 246 AlaGluTrpLeuPro-----GluLeuGlyLeuThrArgGluSer---TyrGluVal 261
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 242 GCCCAAGCTTGCCCGGTGCGGTGCGACTTGCCCTTCCACCGCAAGTCTGCCCGCATAT 301
OY 262 HisAlaValPheValProGluMetAlaIleAspLeuGluAlaGluValLysArgLeuArg 281
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 302 CACGCGATTGGATTCCGGAACCTACAGCTAGACTGTCGCGAGTCCGCGGCTTTGAAG 361
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
OY 282 GluValMetAspLysValAspCysValAsnIlePheValSerGluGlyValGlyGlu 301
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```


Db 362 AAGATGATGACCAAGTCGGTTCGTCACATCTTTTCGCGGACGACGGCGTGCAG 421
Qy 302 AAlaIleValAlaGluMetGlnAlaValysGlyGlnGluValProArgAspAlaPheGlyHis 321
Db 422 GAATCGTCCGCGGACATGAGGCGCATGTGACGCGCGTGCAGCATGCTTTGGACAC 481
Qy 322 IleValLeuAspAlaValAsnProGlyLysTrpPheGlyGlnPhe--AlaGlnMet 340
Db 482 GTCACCTTGCGCTAAATTCAMCCCGGACGACTCTCCACACATTTGCGCGGACAAAT 541
Qy 341 IleGlyAlaGluLysTrpLeuVal 348
Db 542 ATTCGTCGCGAAGACCATTTG 565

RESULT 15
CD375439 677 bp mRNA linear EST 31-MAY-2003
LOCUS PTM000602 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
DEFINITION tricornutum cDNA 5', mRNA sequence.
ACCESSION CD375439
VERSION CD375439.1 GI:31251053
KEYWORDS EST.
SOURCE Phaeodactylum tricornutum
ORGANISM Phaeodactylum tricornutum
Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
1 (bases 1 to 677)
AUTHORS Scala,S., Caroti,N., Falciatore,A., Chiusano,M.L. and Bowler,C.
TITLE Genome properties of the diatom Phaeodactylum tricornutum
JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
MEDLINE 2211123
PUBMED 12114555

COMMENT
Contact: Bowler C
Laboratory of Molecular Plant Biology
Stazione Zoologica 'Anton Dohrn'
Villa Comunale, I-80121, Napoli, Italy
Tel: 39 081 583 3268/3211
Fax: 39 081 764 1355
Email: chris@alpha.szn.it
Diatom EST Database(<http://aves.thagen.szn.bowler.com>)
Seq primer: T3 backward
POLY-A-No.

FEATURES
source
1..677
/location/Qualifiers
/organism="Phaeodactylum tricornutum"
/mol_type="mRNA"
/db_xref="taxon:2850"
/cell_line="CMP632"
/clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
/note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
Xho I"

ORIGIN
Alignment Scores:
Pred. No.: 3,7e-48 Length: 677
Score: 485.50 Matches: 93
Percent Similarity: 73.264 Conservative: 33
Best Local Similarity: 54.074 Mismatches: 44
Query Match: 21.264 Indels: 2
DB: 6 Gaps: 1

US-10-701-200-6 (1-437) x CD375439 (1-677)

Qy 256 ArgGluSerTrpGluValHisAlaValPheValProGluMetAlaIleAspLeuGluAla 275
Db 51 AAGTCGCGCGCGCATATTCACGCGCATTTGGATTCCGAACCTCAAGCTTAGCTTGGCGCC 110
Qy 276 GluAlaLysArgLeuArgGluValMetAspLysValAspCysValAsnIlePheValSer 295
Db 111 GAGTCGCGCGCGCTTGAAGAATCATGAGCAAGTCGCTTGCCTCAACATCTTTTCGCGC 170
Qy 296 GluGlyAlaGlyValGluAlaIleValAlaGluMetGlnAlaLysGlyGlnGluValPro 315

Db 171 GAGGCGACCGGCGCTGCAGAAATCGTCGCGACATGAGAGCCAAAGGTGAAGCCGTGCCG 230
Qy 316 ArgAspAlaPheGlyHisIleLysLeuAspAlaValAsnProGlyLysTrpPheGlyGlu 335
Db 231 CGGATGCTCTTTGGACACGTCACGTTGGCTAAATCATATCCGGACAGTACTTCCCAA 250
Qy 336 GlnPheAlaGlnMetIleGlyAlaGluLysTrpLeuValGlnLysSerGlyTyrPheAla 355
Db 291 CACTTGGCGGACAAATATTGGTGGCGAAGAACCATTTGCGAAGATCGGGAATCTTGGCC 350
Qy 356 ArgAlaSerAlaSerAsnValAspAspMetArgLeuIleLysSerCysAlaAspLeuAla 375
Db 351 CGTTCCGCGCGCGCAACGATTTGATCCGCAACTCATCGGCGCTGTGCGAGGCTGCC 410
Qy 376 ValGluCysAlaPheArgArgGluSerGlyValIleGlyHisAspGluAspAsnGlyAsn 395
Db 411 GTGCGCGCGCGCTATTGACGACATTCGCGATGATGGACAGATGAAGCAAAACCCAC 470
Qy 396 Val--LeuArgAlaIleGluPheProArgIleLysGlyGlyLys-ProPheAsnIleAs 414
Db 471 AGCGCCATTCGAGCGCATTTGAATTCAGTCCGATCAAGGTGGCAAAACCCCTTGATATTTC 530
Qy 414 pThrAspTrpPheAsnSerMetLeuSerGluIle 425
Db 531 TCAGGAATGGGTCCAAACAAATGCTCAANGAATAAT 564

Search completed: March 1, 2005, 18:27:33
Job time : 4383 secs

THIS PAGE BLANK (USPTO)